# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-028937

(43)Date of publication of application : 31.01.1995

(51)Int.Cl. G06K 9/34

(21)Application number : 05-168253

(22)Date of filing : 07.07.1993

(71)Applicant : FUJITSU LTD

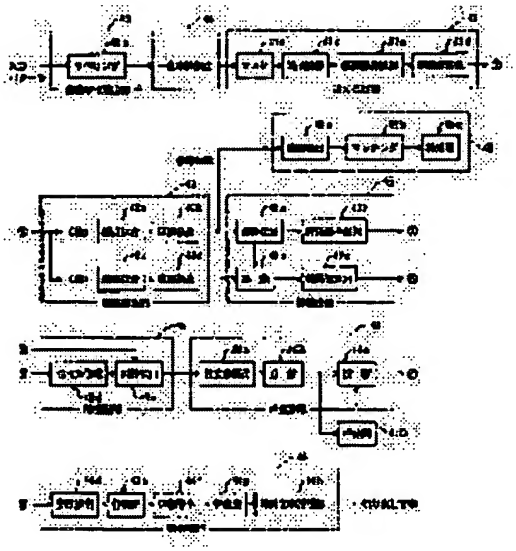
(72)Inventor : NAOI SATOSHI
ASAKAWA ATSUKO
YABUKI MASANORI
HOTSUTA YOSHINOBU

## (54) IMAGE EXTRACTING DEVICE

(57)Abstract:

PURPOSE: To provide an image extracting device which accurately extract the frames out of the images including the characters, etc., touching the frames and restores characters, etc., in high quality.

CONSTITUTION: An image extracting device consists of a coupling pattern extracting means 40 which extracts a partial pattern where the picture elements are coupled to each other out of an image comprising either a block frame of a table form including a single character frame or a free format frame and characters, graphics or symbols, a single character frame extracting means 46 which extracts a single character frame out of the image based on the extracted partial pattern, (the straight line extracting means 41 and 42 which extract the straight lines out of the partial pattern with a single character frame removed), a frame detecting means 43 which detects the straight lines that construct a frame, and a frame separating means 44 which segments the characters, graphics or symbols by separating the detected straight lines from the partial pattern.



LEGAL STATUS

| | |
|---|---|
| [Date of request for examination] | 05.09.1997 |
| [Date of sending the examiner's decision of rejection] | 26.10.1999 |
| [Kind of final disposal of application other than the examiner's decision of rejection or application converted registration] | |
| [Date of final disposal for application] | |
| [Patent number] | 3042945 |
| [Date of registration] | 10.03.2000 |
| [Number of appeal against examiner's decision of rejection] | 11-18548 |
| [Date of requesting appeal against examiner's decision of rejection] | 24.11.1999 |
| [Date of extinction of right] | |

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

**JPO and NCIPI are not responsible for any damages caused by the use of this translation.**

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

## CLAIMS

[Claim(s)]
[Claim 1] A connection pattern extract means to extract the partial pattern with which the pixel and the pixel are connected from the image which consists of a block frame or a free-formatting frame, and the alphabetic character, graphic form or notation of the tabular format containing a single-character frame (40), A single-character frame extract means to extract a single-character frame from said image based on the partial pattern extracted by this connection pattern extract means (40) (46), A straight-line extract means to detect a straight line from the partial pattern which was extracted by this connection pattern extract means (40), and was removed by this single-character frame extract means (46) in the single-character frame (41 42), A frame detection means to detect the straight line which constitutes a frame from a straight line detected by this straight-line extract means (41 42) (43), By separating the straight line detected by this frame detection means (43) from a partial pattern, an alphabetic character, It has the frame separation means (44) which starts a graphic form or a notation. This straight-line extract means (41 42) For every partial pattern which was extracted by this connection pattern extract means (40), and was removed in the single-character frame by this single-character frame extract means (46) The means which scans within two kinds of oblong and longwise masks to the whole image of the partial pattern removed in the single-character frame, If a means to compute the rate that the pattern in said mask occupies, and the computed value are larger than a predetermined value, while considering that all the inside of the mask is patterns An image extractor including a means to extract a component in every direction by deleting the pattern in a mask if the computed value is below a predetermined value.
[Claim 2] Said straight-line extract means (41 42) is an image extractor of claim 1 which summarizes those lines or trains, creates the rectangle range, and includes the means which makes the center line of the rectangle range a processing result, when two or more lines or trains continue and said rate becomes larger than a predetermined value.
[Claim 3] Said straight-line extract means (41 42) is claim 1 or the image extractor of 2 including a means to give a lap respectively to the range which performs mask processing.
[Claim 4] A connection pattern extract means to extract the partial pattern with which the pixel and the pixel are connected from the image which consists of a block frame or a free-formatting frame, and the alphabetic character, graphic form or notation of the tabular format containing a single-character frame (40), A single-character frame extract means to extract a single-character frame from said image based on the partial pattern extracted by this connection pattern extract means (40) (46), A straight-line extract means to detect a straight line from the partial pattern which was extracted by this connection pattern extract means (40), and was removed by this single-character frame extract means (46) in the single-character frame (41 42), A frame detection means to detect the straight line which constitutes a frame from a straight line detected by this straight-line extract means (41 42) (43), It has the frame separation means (44) which starts an alphabetic character, a graphic form, or a notation by separating the straight line detected by this frame detection means (43) from a partial pattern. This frame detection means (43) A means to detect the thing of the die length beyond a predetermined value as a transversal frame among the candidates of the straight line which constitutes the transversal frame detected with this straight-line

detection means (41 42), An image extractor including a means to detect the block frame of a tabular format when a transversal frame is three or more while detecting a horizontal party's block frame or free-formatting frame, when the number of the transversal frames which adjoin based on the detected transversal frame is two.
[Claim 5] Said frame detection means (43) is based on the candidate and said detected transversal frame of the straight line which constitutes the door post detected with said straight-line detection means (41 42). It is the image extractor of claim 5 which includes the means which others consider that is a part of alphabetic character, and is excluded by a means by which retrieval investigates whether even the up-and-down transversal frame with which the straight line of the candidate of a door post was detected has reached, or it has not broken off by the middle, and the vertical line to which the upper and lower sides have reached the transversal frame as a result of retrieval considering as the candidate of a door post.
[Claim 6] Said frame detection means (43) is claim 4 or the image extractor of 5 including the means which shall be a door post about the vertical line which was not excepted from the candidate of a door post by excepting the vertical line which forms remarkably different spacing from others among the vertical lines within the limits inserted into two detected ****** transversal frames from the candidate of a door post, when the object of processing is the block frame of a regular tabular format.
[Claim 7] For said frame detection means (43), the straight line which exceeds predetermined die length in the case of a frame extract is claims 4 and 5 or the image extractor of 6 including a means to divide into the unit of fixed die length.
[Claim 8] Said frame detection means (43) is the image extractor including a means to change the number and location of the frame eliminated by to which location in a free-formatting frame the divided part corresponds of claim 7.
[Claim 9] Said frame detection means (43) is claims 4, 5, 6, and 7 or the image extractor of 8 which separates [ each ] from the left one part surrounded by the closing line in the four way type, goes, and includes a means to treat each leaver section as a single-character frame when the ratio of the die length of striping and a vertical line is below a predetermined value.
[Claim 10] A connection pattern extract means to extract the partial pattern with which the pixel and the pixel are connected from the image which consists of a block frame or a free-formatting frame, and the alphabetic character, graphic form or notation of the tabular format containing a single-character frame (40), A single-character frame extract means to extract a single-character frame from said image based on the partial pattern extracted by this connection pattern extract means (40) (46), A straight-line extract means to detect a straight line from the partial pattern which was extracted by this connection pattern extract means (40), and was removed by this single-character frame extract means (46) in the single-character frame (41 42), A frame detection means to detect the straight line which constitutes a frame from a straight line detected by this straight-line extract means (41 42) (43), The frame separation means which starts an alphabetic character, a graphic form, or a notation by separating the straight line detected by this frame detection means (43) from a partial pattern (44), The image extractor which has a frame extract / removal means (45) to perform frame extract and removal processing about the partial pattern in which the transversal frame detection in this straight-line extract means (41 42) failed.
[Claim 11] The means which unifies segments also including the segment to which said frame extract / removal means (45) has broken off on the way to the partial pattern in which the transversal frame retrieval in said straight-line extract means (41 42) failed, When striping after integration has the die length beyond a predetermined value as compared with the magnitude of a partial pattern, while considering that the striping is a transversal frame The means which makes what is attained to up-and-down striping among the vertical lines between two ****** striping the candidate of a door post, Thus, the single-character frame range is extracted by being in the same image as the rectangle range formed of the candidate of the obtained frame and a frame, and taking matching with a single-character frame [ finishing / an extract ] with said single-character frame extract means (46). The image extractor including a means to perform extract of a frame, and processing of removal like the case of a single-character frame to the one letter-face each range of claim 10.
[Claim 12] A connection pattern extract means to extract the partial pattern with which the pixel and the

pixel are connected from the image which consists of a block frame or a free-formatting frame, and the alphabetic character, graphic form or notation of the tabular format containing a single-character frame (40), A single-character frame extract means to extract a single-character frame from said image based on the partial pattern extracted by this connection pattern extract means (40) (46), A straight-line extract means to detect a straight line from the partial pattern which was extracted by this connection pattern extract means (40), and was removed by this single-character frame extract means (46) in the single-character frame (41 42), A frame detection means to detect the straight line which constitutes a frame from a straight line detected by this straight-line extract means (41 42) (43), By separating the straight line detected by this frame detection means (43) from a partial pattern, an alphabetic character, It is an image extractor including the means made into the start point of retrieval of the thinnest part within the rectangle segment judged that it has the frame separation means (44) which starts a graphic form or a notation, and it is a frame in case this straight-line extract means (41 42) searches for a segment, in order to detect a straight line from a partial pattern.

[Claim 13] In the image extractor which extracts a frame from the image which consists of the frame which each rectangle separated with the rectangle, an alphabetic character, a graphic form, or a notation A connection pattern extract means to extract the partial pattern with which the pixel and the pixel are connected from the pattern which constitutes an image (21), A frame extract means to extract a frame based on each extracted partial pattern (22), An intersection calculation means to compute the intersection of an alphabetic character and a frame based on the partial pattern and frame which were extracted (25-1, 25-2), The intersection matching means which the criterion of the distance between the alphabetic character segments whose frames are pinched according to the line breadth of a frame, and the continuity of an inclination is changed accommodative, and matches an intersection based on the continuity of said distance and an inclination (55 61), The image extractor which has a within the limit alphabetic character extract means (56) to extract an alphabetic character segment within the limit based on the matched intersection.

[Claim 14] Said intersection calculation means (25-1, 25-2) is an image extractor of claim 13 which includes a means to search for an alphabetic character segment in the direction of a core of a frame from an intersection side, and a means to judge whether the searched intersection of a result within the limit is extension of an alphabetic character segment, about all the intersections of the frame and alphabetic character which were computed.

[Claim 15] For the intersection, said intersection calculation means (25-1, 25-2) is the image extractor including a means to judge that it is not the intersection of a frame and an alphabetic character segment of claim 14, when an intersection within the limit in the process in which it searches for an alphabetic character segment in the direction of a core of a frame from the intersection side of a frame and an alphabetic character segment stops existing.

[Claim 16] Said intersection calculation means (25-1, 25-2) is claim 14 or the image extractor of 15 including a means to ask for the area of the alphabetic character segment for which is the process in which it searches for an alphabetic character segment in the direction of a core of a frame from the intersection side of a frame and an alphabetic character segment, and it searched in the direction of a core of a frame from the intersection side.

[Claim 17] Said intersection calculation means (25-1, 25-2) is an image extractor given in any 1 term among claims 13-16 including a means to ask for the area of the alphabetic character segment for which the direction of a core of the intersection side of a frame and an alphabetic character segment to a frame is the process in which it searches for an alphabetic character segment to hard flow, and it searched from the intersection side to said hard flow.

[Claim 18] Said intersection calculation means (25-1, 25-2) is the image extractor including a means to judge the alphabetic character segment to be a noise when the sum of the area of each alphabetic character segment for which it searched to said direction of a core and hard flow is smaller than a predetermined threshold of claim 17, when an intersection outside the limit in the process in which it searches for an alphabetic character segment from the intersection side of a frame and an alphabetic character segment to the direction of a core and hard flow of a frame stops existing.

[Claim 19] When the contact frequency where compute the contact frequency of an alphabetic character and a frame from the degree of contact in a frame, the width of character and an alphabetic character, and a frame, and the alphabetic character and the frame do not touch in [ a character width is larger than frame width, and ] comparison is small, a character width and frame width -- abbreviation -- the inside of claims 13-18 which have further a contact frequency calculation means (51) to be the same or to distinguish the case where contact frequency with comparatively much contact in an alphabetic character and a frame is large -- an image extractor given in any 1 term.
[Claim 20] When the contact frequency which said contact frequency calculation means (51) computed is large, said intersection matching means (55 61) An intersection outside the limit, The image extractor of claim 19 which searches for the directivity of an alphabetic character segment from three kinds of intersections of the intersection of an alphabetic character and a frame, and an intersection within the limit, and includes the means which matches the intersection of an alphabetic character and a frame based on the conditions and said directivity of a continuity of the alphabetic character segment based on the distance and the direction of [ between the alphabetic character segments in an intersection within the limit ].
[Claim 21] said intersection matching means (55 61) -- abbreviation -- the image extractor of claim 20 which matches intersections with the same directivity.
[Claim 22] Claim 20 or 21 image extractors which have further a within the limit alphabetic character extract means (56) to extract the range surrounded within the intersection matched with the means which matches the intersection of an alphabetic character and a frame based on said three kinds of intersections, and the limit as an alphabetic character component.
[Claim 23] Said intersection matching means (55 61) is an image extractor given in any 1 term among claims 19-22 which search for the directivity of an alphabetic character segment from two kinds of intersections of an intersection outside the limit and the intersection of an alphabetic character and a frame, and include the means which matches the intersection of an alphabetic character and a frame based on said directivity when the contact frequency which said contact frequency calculation means (51) computed is small.
[Claim 24] said intersection matching means (55 61) -- abbreviation -- the image extractor of claim 23 which matches intersections with the same directivity.
[Claim 25] Claim 23 or 24 image extractors which have further a within the limit alphabetic character interpolation means (32) to interpolate an alphabetic character field within the limit by connecting between the intersections matched with the means which matches the intersection of an alphabetic character and a frame based on said two kinds of intersections.
[Claim 26] A character string extract means to extract a character string after matching of the intersection by said intersection matching means (55 61) (12-1, 12-2), The integrated means which unifies an alphabetic character based on the information about an average character pitch and an average character size at least (16-1, 16-2), It is an image extractor given in any 1 term among claims 23-25 which have further a connective check means (57-1, 57-2) to compute the range of the frame which checks and re-interpolates the connectivity of an alphabetic character component based on the location gap with the location of a frame and the location of an alphabetic character segment which were extracted.
[Claim 27] Said intersection matching means (55 61) is an image extractor of claim 26 which performs re-matching of an intersection which recovers connectivity when the connectivity of an alphabetic character component cannot be checked with a connective check means (57-1, 57-2).
[Claim 28] It is an image extractor given in any 1 term among claims 23-25 which have further the integrated means (16-1, 16-2, 16-3) which unifies an alphabetic character based on the information about an average character pitch and an average character size at least, and a re-interpolation judging means (63) to make said integration perform according to the existence of the re-interpolated character pattern.

---

[Translation done.]

* NOTICES *

JPO and NCIPI are not responsible for any
damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

## DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Industrial Application] Especially this invention relates to the image extractor for extracting an alphabetic character, a graphic form, etc. which contacted the letter face, the ruled line, etc. in handwritten character recognition systems, such as OCR, with respect to an image extractor.
[0002] The need of a handwritten character recognition system is increasing as an I/O device for handwriting. In order to realize the high recognition rate of a discrete character in such a handwritten character recognition system, it is important that logging processing of the alphabetic character which is the preceding paragraph story of recognition is performed correctly.
[0003] There is a document beforehand specified in the location which writes alphabetic characters, such as a document, as a document set as the object of recognition. In such a document, the frame which specified the location which writes an alphabetic character is written by a color and concentration with same ruled line and alphabetic characters, such as a black frame instead of a drop out color. Therefore, in the comparatively high recognition rate, when finely written to within the limits specified by an alphabetic character, although automatic recognition was possible, when the handwriting alphabetic character contacted the frame or ruled line which shows the appointed range across the appointed range or disturbed, the problem that a recognition rate fell remarkably had arisen.
[0004] This invention relates to the image extractor for extracting an alphabetic character, a graphic form, a notation, etc. from the alphabetic character which is touching the ruled line, the frame, etc., a graphic form, a notation, etc. correctly to the above-mentioned problem. That is, this invention relates to the image extractor which can be applied in case only a graphic form is cut down from the pattern with which the graphic form in large semantics lapped with the straight line like separation of the contact part of the ruled line in logging of the alphabetic character in not only the character reader for handwriting but printer graphic recognition equipment or drawing recognition equipment and a notation, and an image, and a body, a graphic form and an alphabetic character.
[0005]
[Description of the Prior Art] These people proposed the **** image sampling procedure previously shown in drawing 97 by Japanese Patent Application No. No. 103257 [ five to ]. The input configuration set as the object of processing shall be the binary image by which amendment of an extreme inclination and rotation, removal of a noise, stopgap of a blur, etc. were pretreated beforehand, for example, shall remove a frame from a black-framed document. That is, about the letter face, the handwriting alphabetic character is written to the block frame which there are two or more block frames of long and slender one line horizontally, and the size, location, and inclination of the frame do not understand, and only a frame is removed, even if an alphabetic character contacts a frame or has overflowed the frame.
[0006] An image sampling procedure consists of the connection pattern extract section 40, the segment detecting element 41, the straight-line detecting element 42, a frame detecting element 43, and the frame separation section 44, as shown in drawing 97 . The connection pattern extract section 40 extracts the partial pattern with which the pixel and the pixel are connected from the input configuration which

consists of straight-line parts, such as a frame or a ruled line, an alphabetic character, a graphic form, or a notation. The segment detecting element 41 detects a part of segment or straight line by fixed die length for every connection pattern by contiguity projection including the thinning section. The straight-line detecting element 42 unifies two or more acquired segments or linear [ a part of ], and detects a long straight line. The frame detecting element 43 extracts the straight line which constitutes a letter face with spacing of two or more obtained straight lines etc. The frame separation section 44 computes the width of face of each frame by dividing a block frame into the frame for every single character by the intersection of a frame and a frame, and separates a letter face from a connection pattern according to the width of face.

[0007] On the other hand, when the frames which should be removed from the target input configuration are frames other than the above, the frame extract approach which is proposed in JP,62-212888,A or JP,3-126186,A is also proposed. According to these frame extract approaches, format information, such as a location and size, and the information about an inclination are beforehand inputted as document data, and are stored. Removal of a frame is performed based on such stored information.

[0008]

[Problem(s) to be Solved by the Invention] By the way, in the image sampling procedure which this application proposed previously, there are the following troubles and there is still room of an improvement. When the frame made [ 1st ] into an object was except a single-character frame and horizontal one line's block frame, in the case of the block frame of a tabular format, or the free-formatting frame, frame extract processing was not able to be performed. Since the segment detecting element 41 performed thinning to the 2nd including the thinning section, the processing time was very long. Moreover, if thinning is performed, since the linearity of a subject-copy image will be lost, it becomes as [ show / the subject-copy image shown, for example in drawing 98 (a) / by thinning / drawing 98 (b) ], and the extract of a frame has become difficult. When one alphabetic character had connected the single-character frame of ****** plurality to the 3rd mutually and it turned out that the middle of a straight line has broken off by retrieval, it did not process, after considering that they are not a block frame or a single-character frame, either. Since it searched to it and a frame extract was carried out to it when removing a single-character frame to the 4th, retrieval went wrong and a frame extract might be unable to be performed.

[0009] The concrete example which shows the 4th trouble of the above to drawing 99 explains. the case where the start point of retrieval is A -- retrieval of a partial pattern -- this example -- the direction of a top to the bottom -- and since it progresses to right and left only several predetermined pixel minutes corresponding to the width of face of a frame when it cannot progress in the direction of [ lower ], retrieval is performed good, as the drawing 99 Nakaya mark shows. However, since a partial pattern does not exist in the direction of [ lower ] although retrieval of a partial pattern cannot be advanced in the direction of [ lower ] and predetermined progresses to right and left several pixel minutes when the start point of retrieval is B, as the drawing 99 Nakaya mark shows, retrieval will go wrong. Therefore, in the case of this latter, a frame extract cannot be performed.

[0010] On the other hand, the extract of a frame is possible even if the target frame is a block frame of a tabular format etc. by the frame extract approach which is proposed in JP,62-212888,A or JP,3-126186,A. On the other hand, since it was the requisite to input beforehand format information, such as a location and size, and the information about an inclination as document data, and to store them, there was the 5th trouble that processing tends to be influenced of the irregularity of the letter face itself or few inclinations. That is, it considered that the letter face protruded from the position was an alphabetic character, and it might not extract at the time of the extract of a letter face, and the alphabetic character part might be conversely extracted as a frame. Furthermore, the input of the information about each frame by the user took time amount very much, and there was also a problem that the burden to a user was large in it.

[0011] This invention is made in order to improve the above-mentioned trouble, and straight-line parts, such as a frame and a ruled line, are correctly extracted and separated from the alphabetic character which is touching straight-line parts, such as a frame and a ruled line, a graphic form, a notation, etc. By

starting an alphabetic character, a graphic form, a notation, etc. correctly, the alphabetic character in contact with straight-line parts, such as a frame and a ruled line, a graphic form, a notation, etc. can be restored by high definition, and it aims at offering the image extractor which can raise these recognition rates remarkably.

[0012]

[Means for Solving the Problem] The block frame or free-formatting frame of a tabular format with which the 2nd technical problem of the above contains a single-character frame according to claim 1, A connection pattern extract means 40 to extract the partial pattern with which the pixel and the pixel are connected from the image which consists of an alphabetic character, a graphic form, or a notation, A single-character frame extract means 46 to extract a single-character frame from said image based on the partial pattern extracted by this connection pattern extract means 40, Straight-line extract means 41 and 42 to detect a straight line from the partial pattern which was extracted by this connection pattern extract means 40, and was removed by this single-character frame extract means 46 in the single-character frame, A frame detection means 43 to detect the straight line which constitutes a frame from a straight line detected by these straight-line extract means 41 and 42, It has the frame separation means 44 which starts an alphabetic character, a graphic form, or a notation by separating the straight line detected by this frame detection means 43 from a partial pattern. These straight-line extract means 41 and 42 For every partial pattern which was extracted by this connection pattern extract means 40, and was removed in the single-character frame by this single-character frame extract means 46 The means which scans within two kinds of oblong and longwise masks to the whole image of the partial pattern removed in the single-character frame, If a means to compute the rate that the pattern in said mask occupies, and the computed value are larger than a predetermined value, while considering that all the inside of the mask is patterns If the computed value is below a predetermined value, it will be attained by deleting the pattern in a mask by the image extractor including a means to extract a component in every direction.

[0013] The 1st and 5th technical problems of the above are claim 4 publications. The block frame or free-formatting frame containing a single-character frame of a tabular format, A connection pattern extract means 40 to extract the partial pattern with which the pixel and the pixel are connected from the image which consists of an alphabetic character, a graphic form, or a notation, A single-character frame extract means 46 to extract a single-character frame from said image based on the partial pattern extracted by this connection pattern extract means 40, Straight-line extract means 41 and 42 to detect a straight line from the partial pattern which was extracted by this connection pattern extract means 40, and was removed by this single-character frame extract means 46 in the single-character frame, A frame detection means 43 to detect the straight line which constitutes a frame from a straight line detected by these straight-line extract means 41 and 42, It has the frame separation means 44 which starts an alphabetic character, a graphic form, or a notation by separating the straight line detected by this frame detection means 43 from a partial pattern. This frame detection means 43 A means to detect the thing of the die length beyond a predetermined value as a transversal frame among the candidates of the straight line which constitutes the transversal frame detected with these straight-line detection means 41 and 42, When the number of the transversal frames which adjoin based on the detected transversal frame is two, while detecting a horizontal party's block frame or free-formatting frame, it is attained by the image extractor ****** extractor including a means to detect the block frame of a tabular format when a transversal frame is three or more.

[0014] The block frame or free-formatting frame of a tabular format with which the 3rd technical problem of the above contains a single-character frame according to claim 9, A connection pattern extract means 40 to extract the partial pattern with which the pixel and the pixel are connected from the image which consists of an alphabetic character, a graphic form, or a notation, A single-character frame extract means 46 to extract a single-character frame from said image based on the partial pattern extracted by this connection pattern extract means 40, Straight-line extract means 41 and 42 to detect a straight line from the partial pattern which was extracted by this connection pattern extract means 40, and was removed by this single-character frame extract means 46 in the single-character frame, A frame detection means 43 to detect the straight line which constitutes a frame from a straight line detected by

these straight-line extract means 41 and 42, The frame separation means 44 which starts an alphabetic character, a graphic form, or a notation by separating the straight line detected by this frame detection means 43 from a partial pattern, It is attained by the image extractor which has a frame extract / removal means 45 to perform frame extract and removal processing about the partial pattern in which the transversal frame detection in these straight-line extract means 41 and 42 failed.
[0015] The block frame or free-formatting frame of a tabular format with which the 4th technical problem of the above contains a single-character frame according to claim 11, A connection pattern extract means 40 to extract the partial pattern with which the pixel and the pixel are connected from the image which consists of an alphabetic character, a graphic form, or a notation, A single-character frame extract means 46 to extract a single-character frame from said image based on the partial pattern extracted by this connection pattern extract means 40, Straight-line extract means 41 and 42 to detect a straight line from the partial pattern which was extracted by this connection pattern extract means 40, and was removed by this single-character frame extract means 46 in the single-character frame, A frame detection means 43 to detect the straight line which constitutes a frame from a straight line detected by these straight-line extract means 41 and 42, It has the frame separation means 44 which starts an alphabetic character, a graphic form, or a notation by separating the straight line detected by this frame detection means 43 from a partial pattern. These straight-line extract means 41 and 42 It is attained by the image extractor including the means made into the start point of retrieval of the thinnest part within the rectangle segment judged to be a frame in case it searches for a segment, in order to detect a straight line from a partial pattern.
[0016] In the image extractor which extracts a frame from the image with which invention of claim 12 consists of the frame which each rectangle separated with the rectangle, an alphabetic character, a graphic form, or a notation A connection pattern extract means 21 to extract the partial pattern with which the pixel and the pixel are connected from the pattern which constitutes an image, A frame extract means 22 to extract a frame based on each extracted partial pattern, The intersection calculation means 25-1 and 25-2 which compute the intersection of an alphabetic character and a frame based on the partial pattern and frame which were extracted, The intersection matching means 55 and 61 which the criterion of the distance between the alphabetic character segments whose frames are pinched according to the line breadth of a frame, and the continuity of an inclination is changed accommodative, and match an intersection based on the continuity of said distance and an inclination, It is the image extractor which has a within the limit alphabetic character extract means 56 to extract an alphabetic character segment within the limit based on the matched intersection.
[0017]
[Function] A segment can be detected by the comparatively short processing time, without according to invention according to claim 1, losing the linearity of a subject-copy image, since mask processing is performed.
[0018] According to invention according to claim 4, the block frame and free-formatting frame of a tabular format can be extracted, without needing format information.
[0019] According to invention according to claim 9, even if a way piece is found in a straight line and processing finishes with it, by performing matching processing, the single-character range can be extracted and a frame can be extracted.
[0020] According to invention of claim 11, the success percentage of retrieval of a segment can be improved.
[0021] When frame width is larger than the character width of the alphabetic character in contact with a frame according to invention of claim 12, even if it is the case where the frame inclines etc., matching of an intersection etc. is performed more to accuracy and the logging precision of an alphabetic character can be improved.
[0022]
[Example] Drawing 1 is drawing showing the 1st example of the image extractor which becomes this invention. The input configuration set as the object of processing shall be the binary image by which amendment of an extreme inclination and rotation, removal of a noise, stopgap of a blur, etc. were

pretreated beforehand. Moreover, in this example, a frame shall be removed from a black-framed document for convenience. That is, the handwriting alphabetic character is written to the frame which size, a location, and an inclination do not understand, and only a frame is removed, even if an alphabetic character contacts a frame or has overflowed the frame. The frame set as the object of processing contains the regular block frame of a tabular format shown in drawing 18 (a), the irregular block frame of a tabular format shown in this drawing (b), and the free-formatting frame shown in this drawing (c).
[0023] In this example, as shown in drawing 1 , the connection pattern extract section 40, the single-character frame extract section 46, the segment detecting element 41, the straight-line detecting element 42, the frame detecting element 43, the frame separation section 44, and frame extract / removal section 45 are formed. The connection pattern extract section 40 extracts the partial pattern with which the pixel and the pixel are connected from the input configuration which consists of straight-line parts, such as a frame or a ruled line, an alphabetic character, a graphic form, or a notation. The single-character frame extract section 46 extracts a single-character frame from a partial pattern. The segment detecting element 41 detects a part of segment or straight line by fixed die length for every connection pattern by contiguity projection to the pattern removed in the single-character frame. The straight-line detecting element 42 unifies two or more acquired segments or linear [ a part of ], and detects a long straight line. The frame detecting element 43 extracts the straight line which constitutes a letter face with spacing of two or more obtained straight lines etc. The frame separation section 44 computes the width of face of each frame by dividing a block frame into the frame for every single character by the intersection of a frame and a frame, and separates a letter face from a connection pattern according to the width of face. Frame extract / removal section 45 performs frame extract / removal processing later mentioned about the partial pattern in which the transversal frame detection in the straight-line detecting element 42 failed.
[0024] In addition, the processing to the connection pattern integrated section 34 shown in alphabetic character / frame intersection calculation section 25a - Fig. 3737 can be considered. [ which is shown in drawing 36 mentioned later, for example as processing after separating a frame ]
[0025] The connection pattern extract section 40 consists of labeling section 40a. In order to extract each pattern to stability, without being dependent on the relative relation of the location where two or more block frames are arranged, in labeling section 40a, labeling extracts the pattern connected in 8 connection. The partial pattern obtained by this labeling is either among the frames with which a part of alphabetic character which does not touch a frame when the alphabetic character does not touch, and a frame or alphabetic character, and the alphabetic character touch. Then, these partial patterns are distinguished and a frame is extracted. Moreover, since the size of the partial pattern obtained by labeling is needed later, the coordinate of the angle of the rectangle obtained by carrying out rectangle approximation of the partial pattern is computed during processing of labeling.
[0026] In this example, the single-character frame extract section 56 consists of the segment detecting element 141, the straight-line detecting element 142, a frame detecting element 143, and the frame separation section 144, as shown in drawing 2 . The segment detecting element 141 detects a part of segment or straight line by fixed die length for every connection pattern by contiguity projection. The straight-line detecting element 142 unifies two or more acquired segments or linear [ a part of ], and detects a long straight line. The frame detecting element 143 extracts the straight line which constitutes a letter face with spacing of two or more obtained straight lines etc. The frame separation section 144 computes the width of face of each frame by dividing a block frame into the frame for every single character by the intersection of a frame and a frame, and separates a single-character frame from a connection pattern according to the width of face. The same sign is substantially given to the same part with drawing 1 among drawing 2 .
[0027] The segment detecting element 141 consists of thinning section 41a, contiguity projection section 41b, in-every-direction segment detecting-element 41c, and 41d of the striping part retrieval sections.
[0028] Thinning section 41a performs thinning for every partial pattern obtained by labeling. This thinning makes the size of line breadth the same, for making easy to extract the long straight line which exists only in a frame, it carries out for accumulating and a well-known technique can be used for it as

the thinning itself. In addition, the original pattern before thinning of a partial pattern is memorized apart from the pattern by which thinning was carried out, and is used at the time of retrieval of the segment mentioned later, and frame separation.
[0029] Contiguity projection section 41b divides into plurality in every direction the pattern by which thinning was carried out, is division within the limits of each every direction, computes contiguity projection, and detects the segment of a certain fixed die length, or linear [ a part of ] by rectangle approximation. "Contiguity projection" makes a projection value the result of adding the surrounding line or the projection value of a train to the attention line or the projection value of a train. Thereby, a straight line is detectable even if the straight line is covering the multi-line or the train with the inclination. Therefore, even if it is a frame with a big block frame etc. and is the case where the frame leans, by using contiguity projection, it is possible to detect the straight line which constitutes a frame. For example, if the projection value of i lines is set to p (i) as shown in drawing 3 , contiguity projection value P (i) is computable based on the following (1) type. In addition, in drawing 3 , it is j= 1. Moreover, the projection value of a train as well as the projection value of a line is computable.
[0030]
P [ ] (i) =p(i-j) +p(i-j +1)+ ... +p(i)+ ... +p (i+j) (1)
Drawing 4 is drawing which explains more one example of processing of contiguity projection section 41b to a detail, and shows a flow chart in case software realizes processing.
[0031] In drawing 4 , step S301 divides the partial pattern obtained in the connection pattern extract section 40 in two or more parts of a lengthwise direction and a longitudinal direction. Step S302 is division within the limits of each every direction, and computes projection. Step S303 adds a surrounding projection value to each computed projection value. Furthermore, step S304 computes a contiguity projection value based on the above-mentioned (1) formula.
[0032] In-every-direction segment detecting-element 41c makes the part whose ratio of the contiguity projection value and the division length of each every direction which were computed by contiguity projection section 41b is more than a predetermined threshold a location with a linear candidate. Moreover, for example like the part a of drawing 5 , when two or more lines or trains continue and it becomes more than a predetermined threshold, there shall be a straight line in the rectangle range which summarized them. Since rectangle approximation is carried out, a part of detected straight line or straight line calls it a "rectangle segment" henceforth.
[0033] In addition, if not contiguity projection but the usual projection is used, since a long leaning straight line is undetectable, it needs to make [ many ] the number of partitions of a partial pattern, and needs to shorten division length. However, if division length becomes short, many short straight lines which constitute the alphabetic character will also be detected. On the other hand, since the above-mentioned contiguity projection is used in this example, a comparatively long straight line is also detectable.
[0034] Drawing 6 is drawing which explains more one example of processing of in-every-direction segment detecting-element 41c to a detail, and shows a flow chart in case software realizes processing.
[0035] In drawing 6 , it judges whether step S311 is more than a threshold predetermined in the ratio of the contiguity projection value and the division length of each every direction which were computed by contiguity projection section 41b. If the judgment result of step S311 is NO, it will be judged as that in which a segment does not exist at step S312. On the other hand, if the judgment result of step S311 is YES, it will be judged as that in which a segment exists at step S313. In this case, the segment judged to ****** step S314 judges whether it is in contact with that segment that exists up and down. If the judgment result of step S314 is NO, let the segment judged that step S314 exists be a rectangle segment. On the other hand, if the judgment result of step S314 is YES, the segment judged that step S316 exists and its segment which exists up and down will be unified, and step S317 will make a rectangle segment the segment integrated.
[0036] 41d of striping part retrieval sections performs retrieval for confirming whether the amount of striping has run out on the way among the rectangle segments detected by in-every-direction segment detecting-element 41c. The start point of retrieval is taken as the thinnest part in a rectangle segment.

For example, as shown in this drawing, when fixed locations, such as the middle point shown in drawing 7 "with a circle [ white ]", are made into a start point, and a start point is the part of an alphabetic character, possibility that retrieval will go wrong is high, but since the part thin No. 1 shown by the "black dot" has low possibility of being an alphabetic character, retrieval can be ensured.

[0037] Drawing 8 is drawing which explains more one example of processing of 41d of striping part retrieval sections to a detail, and shows a flow chart in case software realizes processing.

[0038] In drawing 8 , step S321 computes a part with the thinnest line breadth among the rectangle segments detected by in-every-direction segment detecting-element 41c. Step S322 starts retrieval right and left by making the computed part with the thinnest line breadth into a start point. Step S323 searches and step S324 judges whether the branch point exists. If the judgment result of step S324 is YES, the branch point will be memorized at step S325. moreover, a ****** [ that step S326 arrived at / the judgment result of step S324 / the edge of a rectangle segment after NO or step S325 ] -- or it judges whether retrieval went wrong. If the judgment result of step S326 is NO, processing will return to step S323.

[0039] On the other hand, when the judgment result of step S326 is YES, it judges whether retrieval was ended at step S327 and the branch point was memorized at step S328. If the judgment result of step S328 is NO, processing will progress to step S333 mentioned later. If the judgment result of step S328 is YES, it will return to the branch point memorized at step S329, and step S330 will be searched. a ****** [ that step S331 arrived at the edge of a rectangle segment ] -- or it judges whether retrieval went wrong. If the judgment result of step S331 is NO, processing will return to step S330. On the other hand, if the judgment result of step S331 is YES, retrieval will be ended at step S232 and it will judge whether the ending point of retrieval arrived at the edge of right and left of a rectangle segment at step S333. If the judgment result of step S333 is NO, a rectangle segment will judge step S334 not to be a part for striping. Moreover, if the judgment result of step S333 is YES, a rectangle segment will judge step S335 to be a part for striping.

[0040] If it returns to explanation of drawing 2 , the straight-line detecting element 142 will consist of segment integrated section 42a, straight-line detecting-element 42b, segment integrated section 42c, and straight-line detecting-element 42b. Segment integrated section 42a and straight-line detecting-element 42b are prepared to a part for striping, and segment integrated section 42c and straight-line detecting-element 42b are prepared to a part for a vertical line. the segments y and z which show the straight-line detecting element 142 to drawing 9 (a) -- like -- on the way -- the rectangle segments which have not come out and broken off -- contact -- or if connected, these rectangle segments will be unified and it will consider as a long straight line. Furthermore, like the segment x shown in drawing 9 (a), and y, if it is the line added in case a vertical distance is the above-mentioned contiguity projection, or less than the number j of trains even if the rectangle segment is not connected mutually, these rectangle segments will be unified and it will consider as a long straight line. It is judged that straight-line detecting-element 42b is the candidate of the straight line which constitutes a transversal frame if the ratio of the die length of the straight line integrated by segment integrated section 42a and the die length of the rectangle which approximates a partial pattern is more than a predetermined threshold. In addition, straight-line detecting-element 42b asks for the slope of a line integrated, i.e., the inclination of a partial pattern, from the slope of a line which connected the coordinate at the upper left of Segment x and the coordinate at the upper right of Segment z which are integrated by segment integrated section 42a, as shown in drawing 9 (b). The inclination of this partial pattern is used in the processing mentioned later. The same processing as above-mentioned segment integrated section 42a and straight-line detecting-element 42b is performed about all the rectangle segments detected by segment integrated section 42c and straight-line detecting-element 42b about a part for a vertical line. That is, segment integrated section 42c unifies a rectangle segment, and detects a straight line, and it is judged that straight-line detecting-element 42b is the candidate of the straight line which carries out rectangle approximation of the detected straight line, and constitutes a door post.

[0041] Drawing 10 is drawing which explains more one example of processing of segment integrated section 42a of the straight-line detecting element 142, and segment integrated section 42c to a detail, and

shows a flow chart in case software realizes processing.
[0042] In drawing 10 , step S341 computes the distance of the detected rectangle segments. It judges whether step S342 is the line added in case the computed distance is contiguity projection, or less than the number of trains. If the judgment result of step S342 is NO, integration of a rectangle segment shall not be performed at step S343. On the other hand, if the judgment result of step S342 is YES, step S344 will unify a rectangle segment.
[0043] If it returns to explanation of drawing 2 again, the frame detecting element 143 will consist of transversal frame detecting-element 43a, retrieval section 43b, line spacing calculation section 43c, 43d of the histogram creation sections, and door-post detecting-element 43e. Transversal frame detecting-element 43a detects a transversal frame from the candidate of the straight line which constitutes the transversal frame detected by straight-line detecting-element 42b of the straight-line detecting element 42. In this example, since it is aimed at horizontal one line and a block frame with the frame of single-character [ every ] regular intervals, let the thing of the maximum outline be a transversal frame in the straight line of the obtained longitudinal direction. Retrieval section 43b searches for a part for a vertical line based on the transversal frame detected by the linear candidate and transversal frame detecting-element 43a which constitute the door post detected by 42d of straight-line detecting elements of the straight-line detecting element 142 in order to detect a door post. Specifically, the straight line of the candidate of a door post investigates by retrieval whether even the transversal frame of the upper and lower sides obtained by transversal frame detecting-element 43a has reached, or it has not broken off by the middle. The retrieval approach makes a start point the thinnest place among rectangle within the limits like the case of striping. As a result of retrieval, the vertical straight line to which the upper and lower sides have reached the transversal frame considers as the candidate of a door post, and others regard as a part of alphabetic character, and exclude. Next, spacing of the vertical straight lines which remained in the candidate of a vertical line is computed by line spacing calculation section 43c. Moreover, 43d of histogram creation sections shall show the computed line spacing and its frequency of occurrence to a histogram, and door-post detecting-element 43e shall be a door post about the vertical straight line which was not excepted from the candidate of a door post by excepting the vertical straight line which forms remarkably different spacing from others from the candidate of a door post.
[0044] Drawing 11 is drawing which explains more one example of processing of retrieval section 43b to a detail, and shows a flow chart in case software realizes processing. The same sign is substantially given to the same step with drawing 8 among this drawing, and the explanation is omitted.
[0045] In drawing 11 , step S342 starts retrieval up and down by making the thinnest part into a start point. Step S346 and step S351 judge whether the up-and-down transversal frame was reached, respectively, or retrieval went wrong. Step S353 judges whether the point ending [ retrieval ] reached the up-and-down transversal frame. If the judgment result of step S353 is NO, it will judge that step S354 does not have the possibility of a door post, and if a judgment result is YES, it will be judged that step S355 has the possibility of a door post.
[0046] Next, the frame separation section 144 of drawing 2 is explained. The frame separation section 144 consists of frame intersection calculation section 44a, single-character range separation section 44b, projection section 44c, 44d of straight-line detecting elements, frame retrieval section 44e, 44f of **** calculation sections, 44g of the frame removal sections, and 44h of the contact alphabetic character interpolation sections. Frame intersection calculation section 44a computes the intersection of a door post and a transversal frame from the door post obtained from the frame detecting element 143, and a transversal frame. In the example shown in drawing 12 , intersections A1, B1, and A2, B-2, and ... are computed. And single-character range separation section 44b dissociates and goes for the condition (B1, B-2, B3, B4), and ... to the range of every [ a single character ] from the left using the computed intersection (A1, A2, A3, A4), as shown in drawing 12 . Consequently, each leaver section becomes being the same as that of a single-character frame. Projection section 44c computes projection of each leaver section, and 44d of straight-line detecting elements detects a straight line from the obtained projection. 44d of straight-line detecting elements calculates each ratio with the rectangle which carried out rectangle approximation and obtained the projection value and each leaver section of each line of

each leaver section, and a train, and, more specifically, they ask for the candidate of the straight line which constitutes a letter face from the ratio. By computing the distance between straight lines out of the candidate of the straight line which constitutes a letter face, and extracting the straight line of the maximum outline which constitutes a letter face, frame retrieval section 44e searches for each letter face, and asks for the location of a letter face. 44f of **** calculation sections asks for the candidate straight line which adjoins the straight line of the maximum outline, and they compute the line breadth of each searched side of a letter face. 44g of frame removal sections removes every one single-character frame based on the line breadth of each side where the straight line of the maximum outline of each side was location [ the straight line ] and computed. 44h of contact alphabetic character interpolation sections interpolates the alphabetic character part which has been missing by removing a letter face. Thereby, a letter face is removed, and an alphabetic character is started a single character every and goes.

[0047] In addition, when the frame leans like drawing 13 , the parts ra, rb, rc, and rd from which each frame is not removed up and down may remain. Then, when the inclination for which it asked by the above-mentioned straight-line detecting element 142 is comparatively large, the removal range of a 44g [ of frame removal sections ] frame may be made [ more / somewhat ].

[0048] Drawing 14 and drawing 15 are drawings explaining an example at the time of starting an alphabetic character by this example. Drawing 14 (a) shows a original pattern and this drawing (b) shows the frame extracted by 44g of frame removal sections of the frame separation section 44 of drawing 2 . Drawing 1515 (a) shows the alphabetic character started by extracting a frame in 44g of frame removal sections of the frame separation section 44, and this drawing (b) shows the alphabetic character interpolated by 44h of contact alphabetic character interpolation sections of the frame separation section 44.

[0049] As explained above, to the single-character frame divided by the single-character [ every ] frame among the extracted partial patterns, a straight line is extracted from each partial pattern, it judges whether the extracted straight line is a frame, it leaves an alphabetic character field, and the closing line of a single-character frame is eliminated.

[0050] In addition, the extract approach of a single-character frame is not limited to the approach of drawing 2 , and may extract a single-character frame by the part which consists of the connection pattern extract section 1 of drawing 29 mentioned later, the straight-line extract section 2 with thick line breadth, the straight-line extract section 3 with thin line breadth, and the frame extract section 4.

[0051] If it returns and explains to drawing 1 , the segment detecting element 41 will consist of mask-processing section 41A, contiguity projection section 41b, in-every-direction segment detecting-element 41c, and 41d of the striping part retrieval sections.

[0052] Mask-processing section 41A performs mask processing explained below for every partial pattern which removed the single-character frame from the partial pattern obtained by labeling. This mask processing scans within two kinds of oblong and longwise masks to the whole image of the original partial pattern removed in the single-character frame, and computes the rate that the pattern in a mask occupies. If the computed value is larger than a predetermined value, it will be considered that all the inside of the mask is patterns. On the other hand, if the computed value is below a predetermined value, a component in every direction will be extracted by deleting the pattern in a mask. If two or more lines or trains continue and the above-mentioned rate becomes larger than a predetermined value, those lines or trains will be summarized, and the rectangle range will be created, and let the center line of the rectangle range be a processing result.

[0053] In addition, a clearance should just give a lap respectively to open or the range which performs mask processing in order to make it there be nothing between the segments acquired at the result of mask processing. Drawing 16 is drawing explaining the case where the mask which specifies the rectangle range whose vertical x width is 1x6 pixels extracts a horizontal component. (a) shows mask "1" - "9" which overlapped mutually among this drawing. Moreover, this drawing (b) shows an example of the image of a original partial pattern, and this drawing (c) shows the horizontal component extracted when mask processing is performed using the mask shown in (a). The pixel of the horizontal component from which the black round mark was extracted is expressed among this drawing.

[0054] Drawing 17 is drawing which explains more one example of processing of mask-processing section 41A to a detail, and shows a flow chart in case software realizes processing.
[0055] In drawing 17 , step S1 scans within two kinds of oblong and longwise masks to the whole image of a original partial pattern. Step S2 computes the area of the pattern in the mask to the rate, i.e., the magnitude of a mask, that the pattern in a mask occupies, and judges whether this rate is larger than a predetermined value. If the decision result of step S2 is NO, it will judge that there is not a vertical component or a horizontal component at step S3, and processing will finish. On the other hand, when the decision result of step S2 is YES, it considers by step S4 that all the inside of a mask is patterns, and let all be a vertical component or a horizontal component. Step S5 judges whether the obtained component is in contact with the up-and-down component or the component on either side. The component obtained at step S6 as the decision result of step S5 is NO is outputted as a processing result.
[0056] When the decision result of a step SS 5 is YES, the rectangle range is created from the component which unified the component which has touched at step S7, and is integrated at step S8. Step S9 computes the center line of the created rectangle range, and step S10 outputs the center line as a processing result.
[0057] In addition, the original partial pattern before mask processing of a partial pattern is memorized apart from the pattern by which mask processing was carried out, and is used at the time of retrieval of the segment mentioned later, and frame separation.
[0058] Contiguity projection section 41b divides into plurality in every direction the pattern by which mask processing was carried out like contiguity projection section 41b shown in drawing 2 , is division within the limits of each every direction, computes contiguity projection, and detects the segment of a certain fixed die length, or linear [ a part of ] by rectangle approximation. Thereby, a straight line is detectable even if the straight line is covering the multi-line or the train with the inclination. Therefore, even if it is a frame with a big block frame etc. and is the case where the frame leans, by using contiguity projection, it is possible to detect the straight line which constitutes a frame.
[0059] In-every-direction segment detecting-element 41c makes the part whose ratio of the contiguity projection value and the division length of each every direction which were computed by contiguity projection section 41b is more than a predetermined threshold a location with a linear candidate like in-every-direction segment detecting-element 41c shown in drawing 2 . Moreover, for example like the part a of drawing 5 , when two or more lines or trains continue and it becomes more than a predetermined threshold, there shall be a straight line in the rectangle range which summarized them. Since rectangle approximation is carried out, a part of detected straight line or straight line calls it a "rectangle segment."

[0060] In addition, if not contiguity projection but the usual projection is used, since a long leaning straight line is undetectable, it needs to make [ many ] the number of partitions of a partial pattern, and needs to shorten division length. However, if division length becomes short, many short straight lines which constitute the alphabetic character will also be detected. On the other hand, since the above-mentioned contiguity projection is used in this example, a comparatively long straight line is also detectable.
[0061] 41d of striping part retrieval sections performs retrieval for confirming whether the amount of striping has run out on the way among the rectangle segments detected by in-every-direction segment detecting-element 41c like 41d of striping part retrieval sections shown in drawing 2 . The start point of retrieval is taken as the thinnest part in a rectangle segment. For example, as shown in this drawing, when fixed locations, such as the middle point shown in drawing 7 "with a circle [ white ]", are made into a start point, and a start point is the part of an alphabetic character, possibility that retrieval will go wrong is high, but since the part thin No. 1 shown by the "black dot" has low possibility of being an alphabetic character, retrieval can be ensured.
[0062] If it returns to explanation of drawing 1 , the straight-line detecting element 42 will consist of segment integrated section 42a, straight-line detecting-element 42b, segment integrated section 42c, and 42d of straight-line detecting elements. Segment integrated section 42a and straight-line detecting-element 42b are prepared to a part for striping, and segment integrated section 42c and straight-line

detecting-element 42b are prepared to a part for a vertical line. the segments y and z which show the straight-line detecting element 42 to said drawing 9 R> 9 (a) carried out -- like -- on the way -- the rectangle segments which have not come out and broken off -- contact -- or if connected, these rectangle segments will be unified and it will consider as a long straight line. Furthermore, like the segment x shown in drawing 9 (a), and y, if it is the line added in case a vertical distance is the above-mentioned contiguity projection, or less than the number j of trains even if the rectangle segment is not connected mutually, these rectangle segments will be unified and it will consider as a long straight line. It is judged that straight-line detecting-element 42b is the candidate of the straight line which constitutes a transversal frame if the ratio of the die length of the straight line integrated by segment integrated section 42a and the die length of the rectangle which approximates a partial pattern is more than a predetermined threshold. In addition, straight-line detecting-element 42b asks for the slope of a line integrated, i.e., the inclination of a partial pattern, from the slope of a line which connected the coordinate at the upper left of Segment x and the coordinate at the upper right of Segment z which are integrated by segment integrated section 42a, as shown in drawing 9 (b). The inclination of this partial pattern is used in the processing mentioned later. The same processing as above-mentioned segment integrated section 42a and straight-line detecting-element 42b is performed about all the rectangle segments detected by segment integrated section 42c and straight-line detecting-element 42b about a part for a vertical line. That is, segment integrated section 42c unifies a rectangle segment, and detects a straight line, and it is judged that straight-line detecting-element 42b is the candidate of the straight line which carries out rectangle approximation of the detected straight line, and constitutes a door post.

[0063] If it returns to explanation of drawing 1 again, the frame detecting element 43 will consist of transversal frame detecting-element 43a, retrieval section 43b, line spacing calculation section 43c, 43d of histogram creation sections, door-post detecting-element 43e, and contiguity transversal frame selection section 43F. Transversal frame detecting-element 43a detects a transversal frame from the candidate of the straight line which constitutes the transversal frame detected by straight-line detecting-element 42b of the straight-line detecting element 42. In this example, since it is aimed at ****** shown in drawing 18 from which the single-character frame is removed, let the thing of the die length beyond a predetermined value be a transversal frame in the straight line of the obtained longitudinal direction as compared with label size. Contiguity transversal frame selection section 43F choose two adjoining transversal frames based on the transversal frame detected by transversal frame detecting-element 43a. For example, when the number of transversal frames is two, it turns out that it is a horizontal party's block frame or free-formatting frame, and when a transversal frame is three or more, it turns out that it is the block frame of a tabular format. Retrieval section 43b searches for a part for a vertical line based on the transversal frame detected by the linear candidate and transversal frame detecting-element 43a which constitute the door post detected by 42d of straight-line detecting elements of the straight-line detecting element 42 in order to detect a door post. Specifically, the straight line of the candidate of a door post investigates by retrieval whether even the transversal frame of the upper and lower sides obtained by transversal frame detecting-element 43a has reached, or it has not broken off by the middle. The retrieval approach makes a start point the thinnest place among rectangle within the limits like the case of striping. As a result of retrieval, the vertical line to which the upper and lower sides have reached the transversal frame considers as the candidate of a door post, and others regard as a part of alphabetic character, and exclude. Next, spacing of the vertical lines which remained in the candidate of a vertical line is computed by line spacing calculation section 43c. 43d of histogram creation sections shows the computed line spacing and its frequency of occurrence to a histogram. Moreover, door-post detecting-element 43e **** which the object of processing shows to drawing 18 (a) -- in the case of the regular block frame of a tabular format It shall be a door post about the vertical line which was not excepted from the candidate of a door post by excepting the vertical line which forms remarkably different spacing from others among the vertical lines within the limits inserted into two ****** transversal frames detected by 43f of contiguity transversal frame selection sections from the candidate of a door post.

[0064] At this example, a transversal frame considers that the frame which they extract when there are

only two door posts, a left end and a right end, is a free-formatting frame by two. However, since it is hard to detect a long straight line in the case of a frame extract when there is an inclination, as shown in drawing 19 , it divides into the unit of fixed die length. Moreover, the number and location of a frame to eliminate are changed by to which location in a free-formatting frame the divided part corresponds. That is, if the divided part supports the left end location of a free-formatting frame, the frame of the left, a top, and the bottom will be eliminated. If the divided part supports the right end location of a free-formatting frame, the frame of the right, a top, and the bottom will be eliminated. Moreover, if the divided part supports the mid-position of a free-formatting frame, only the frame of a top and the bottom will eliminate **.

[0065] On the other hand, when the frame to extract is a block frame of a tabular format, 43f of contiguity transversal frame selection sections chooses two ****** at a time among transversal frames, and they regard it as that with which a party's block frame used as the transversal frame of the upper and lower sides of the transversal frame of each two ****** as shown in drawing 20 has lapped perpendicularly. Therefore, it is possible to perform the same processing as the case of a party [ every ] 1-block frame in this case.

[0066] in addition, **** which the object of processing shows to drawing 18 (b) -- if spacing (distance) of the vertical lines by which door-post detecting-element 43e remained in the candidate of a door post in the case of the irregular block frame of a tabular format is extremely small as compared with a character size, it shall be a door post by excepting the vertical line from the candidate of a door post about the vertical line which was not excepted from the candidate of a door post moreover, **** which the object of processing shows to drawing 18 (a) -- it shall be a door post about the vertical line which was not excepted from the candidate of a door post by excepting the vertical line which forms remarkably different spacing from others among the vertical lines within the limits which were inserted into two ****** transversal frames detected by 43f of contiguity transversal frame selection sections in the case of the regular block frame of a tabular format from the candidate of a door post

[0067] Drawing 21 is drawing which explains more one example of processing of the frame detecting-element 43 whole to a detail, and shows a flow chart in case software realizes processing.

[0068] In drawing 21 , step S81 judges whether the candidate of the straight line which constitutes the transversal frame detected by straight-line detecting-element 42b of the straight-line detecting element 42 has the die length beyond a predetermined value. The candidate of the straight line from which step S82 constitutes a transversal frame as the decision result of step S81 is NO judges that it is not a transversal frame, and finishes processing. On the other hand, when the decision result of step S81 is YES, a transversal frame is detected from the candidate of the straight line which constitutes a transversal frame from a step S83. Step S84 chooses two adjoining transversal frames based on the detected transversal frame. Moreover, step S85 considers that the range inserted into two adjoining selected transversal frames is one block frame.

[0069] Next, in order to detect a door post, step S86 detects a vertical line based on the candidate of the straight line which constitutes the door post detected by 42d of straight-line detecting elements of the straight-line detecting element 42, and step S87 searches for a part for a vertical line based on the door post and transversal frame which were detected. It judges whether even the up-and-down transversal frame with which the straight line of the candidate of a door post was detected has attained step S88. The vertical line of the candidate of the door post detected at step S89 as the decision result of step S88 is NO is excepted from the candidate of a door post, and processing is finished. On the other hand, when the decision result of step S88 is YES, let the vertical line of the candidate of the door post detected at step S91 be the candidate of a door post.

[0070] Step S92 judges whether the object of processing is the block frame of a regular tabular format, or the block frame of an irregular tabular format. When the object of processing is the block frame of a regular tabular format, while computing spacing of the vertical lines which remained in the candidate of a vertical line at step S93, the histogram showing the relation between the computed line spacing and its frequency of occurrence is computed. **** the object of processing indicates step S94 to be to drawing 18 (a) -- it shall be a door post about the vertical line which was not excepted from the candidate of a

door post by excepting the vertical line which forms remarkably different spacing from others among the vertical lines within the limits which were inserted into two detected ****** transversal frames in the case of the regular block frame of a tabular format from the candidate of a door post Step S95 extracts a frame based on the transversal frame and door post which were obtained as mentioned above, and finishes frame extract processing.

[0071] In addition, if it is judged that the object of processing is the block frame of an irregular tabular format in step S92, it will consider that all the vertical lines of the candidate of a door post are door posts at step S96, and processing will progress to step S95 after that.

[0072] Next, the frame separation section 44 of drawing 1 is explained. The frame separation section 44 consists of frame intersection calculation section 44a, separation section 44B, projection section 44c, frame division section 44D, 44d of straight-line detecting elements, frame retrieval section 44e, 44f of **** calculation sections, 44g of the frame removal sections, and 44h of the contact alphabetic character interpolation sections. Frame intersection calculation section 44a computes the intersection of a door post and a transversal frame from the door post obtained from the frame detecting element 43, and a transversal frame. In the example shown in drawing 22 , intersections A1, B1, and A2, B-2, and ... are computed. And as shown in drawing 22 , using the computed intersection (A1, A2, A3, A4), separation section 44B separates from the left one part surrounded by the closing line in the four way type into each condition (B1, B-2, B3, B4), and ..., and goes. Moreover, among drawing 22 , when the ratio of the die length of striping and a vertical line is below a predetermined value like C1 and C2, each leaver section becomes being the same as that of a single-character frame. Projection section 44c computes projection of each leaver section, and 44d of straight-line detecting elements detects a straight line from the obtained projection. 44d of straight-line detecting elements calculates each ratio with the rectangle which carried out rectangle approximation and obtained the projection value and each leaver section of each line of each leaver section, and a train, and, more specifically, they ask for the candidate of the straight line which constitutes a letter face from the ratio. In addition, the calculation approach of a projection value is explained more to a detail with the 2nd example mentioned later. By computing the distance between straight lines out of the candidate of the straight line which constitutes a letter face, and extracting the straight line of the maximum outline which constitutes a letter face, frame retrieval section 44e searches for each letter face, and asks for the location of a letter face. 44f of **** calculation sections asks for the candidate straight line which adjoins the straight line of the maximum outline, and they compute the line breadth of each searched side of a letter face. 44g of frame removal sections removes every one letter face based on the line breadth of each side where the straight line of the maximum outline of each side was location [ the straight line ] and computed. 44h of contact alphabetic character interpolation sections interpolates the alphabetic character part which has been missing by removing a letter face. Thereby, a letter face is removed, and an alphabetic character is started a single character every and goes.

[0073] In addition, among drawing 22 , when the ratio of the die length of striping and a vertical line is larger than a predetermined value like C3, the number and location of a closing line which are eliminated according to the location of each division part as shown in this drawing are changed, and a closing line is eliminated. Specifically, a frame is divided into two or more leaver sections like the case of a free-formatting frame by frame division section 44D. Furthermore, 44h of projection section 44c - contact alphabetic character interpolation sections performs the above-mentioned processing to each leaver section obtained from frame division section 44D.

[0074] Drawing 23 is drawing which explains more one example of processing of the frame separation section 44 whole to a detail, and shows a flow chart in case software realizes processing.

[0075] In drawing 23 , step S101 computes the intersection of a door post and a transversal frame from the door post obtained from the frame detecting element 43, and a transversal frame. Using the computed intersection, step S102 separates [ each ] from the left one part surrounded by the closing line in the four way type, and goes. It judges whether step S103 has the ratio of the die length of striping and a vertical line larger than a predetermined value. When the ratio of the die length of striping and a vertical line is below a predetermined value, the decision result of step S103 is NO, since each leaver

section becomes being the same as that of a single-character frame, an extract and removal of a frame are performed like the case of a single-character frame at step S104, and frame separation processing is finished.

[0076] On the other hand, when the decision result of step S103 is YES, the number and location of a closing line which are eliminated according to the location of each division part are changed, and a closing line is eliminated. That is, the part surrounded by the closing line at step S105 is divided into two or more leaver sections like the case of a free-formatting frame, and the location of a frame is stored for every leaver section at step S106. Moreover, for every leaver section obtained by carrying out in this way, step S107 performs an extract and removal of a frame like the case of a single-character frame, and finishes frame separation processing.

[0077] Next, frame extract / removal section 45 of drawing 1 is explained. Frame extract / removal section 45 consists of straight-line detecting-element 45a, matching processing section 45b, and frame processing section 45c. Straight-line detecting-element 45a unifies segments also including the segment which has broken off on the way to the partial pattern in which the transversal frame retrieval in the straight-line detecting element 42 failed. Matching processing section 45b considers that the striping is a transversal frame, when striping after integration has the die length beyond a predetermined value as compared with the magnitude of a partial pattern. Moreover, let what is attained to up-and-down striping among the vertical lines between two ****** striping be the candidate of a door post. Thus, the single-character frame range is extracted by being in the same image as the rectangle range formed of the candidate of the obtained frame and a frame, and taking matching with a single-character frame [ finishing / an extract ] by the top Norikazu letter-face extract section 46. Frame processing section 45c performs extract of a frame, and processing of removal like the case of a single-character frame to the one letter-face each range obtained from matching processing section 45b.

[0078] Even if the ********** single-character frame which follows, for example, is shown in drawing 24 (a) is the subject-copy image connected in the alphabetic character, an extract and removal of a frame can be performed good. That is, in such a case, since distinction with a single-character frame and a block frame does not stick only in the magnitude of a partial pattern, the processing as the case of a block frame that the segment detecting element 41 is the same is performed. Moreover, the straight-line detecting element 42 is not recognized to be the straight line from which it constitutes a block frame since striping detects having broken off on the way by retrieval for this striping that broke off. For this reason, it moves to frame extract / removal section 45, and processing unifies striping also including the segment which has broken off on the way to the partial pattern with which the transversal frame retrieval in the straight-line detecting element 42 failed in straight-line detecting-element 45a. Moreover, matching processing section 45b performs matching with the single-character frame [ finishing / an extract / already ] like the above, as shown in drawing 24 (b). This becomes separable into the **** single-character frame shown in drawing 24 (c).

[0079] Drawing 25 is drawing which explains more one example of processing of the frame extract / removal section 45 whole to a detail, and shows a flow chart in case software realizes processing.

[0080] In drawing 25 , step S111 searches for segments also including the segment which has broken off on the way to the partial pattern in which the transversal frame retrieval in the straight-line detecting element 42 failed. Step S112 judges whether the segment has broken off or not as a result of retrieval. If the decision result of step S112 is NO, frame detection processing in which it explained with drawing 21 will be continued. On the other hand, when the decision result of step S112 is YES, striping is unified by detecting a straight line at step S113. the part surrounded in a straight line in the four way type with which step S114 extracted the part surrounded in a straight line in the four way type, and step S115 was extracted -- the magnitude of the single-character frame in the same image, and abbreviation -- it judges whether it is the same. If the decision result of step S115 is NO, a four way type will be judged that the part surrounded in a straight line is not a single-character frame at step S116, and processing will be finished. On the other hand, if the decision result of step S115 is YES, the part surrounded in a straight line in the four way type at step S117 will be recognized as single-character range equivalent to a single-character frame, and extract of a frame and processing of removal will be performed like the case of a

single-character frame.
[0081] By the way, when searching retrieval section 43b in the frame detecting element 43 for a segment, failure in retrieval which was explained with drawing 99 may take place. In case retrieval in the step S87 in drawing 21 R> 1 is performed there, the success percentage of retrieval may be improved by considering as the start point of retrieval of the thinnest part within the rectangle segment judged to be a frame as shown in drawing 26 . Although the frame shown in drawing 26 is the same as the frame shown in drawing 99 , in drawing 26 , it turns out that retrieval of a segment is successful.
[0082] Drawing 27 and drawing 28 are drawings explaining an example at the time of starting an alphabetic character by this example. Drawing 27 (a) shows the original pattern obtained from the connection pattern extract section 40 shown in drawing 1 , and this drawing (b) shows the pattern after the single-character frame extract section 46 shown in drawing 1 removes a single-character frame. Moreover, drawing 27 (c) shows the pattern after performing mask processing by mask-processing section 41A in the segment detecting element 41 shown in drawing 1 to the pattern of this drawing (b). Drawing 28 (a) shows the frame extracted in 44g of frame removal sections in the frame separation section 44 shown in drawing 1 based on the pattern shown in drawing 27 (c), and drawing 28 R> 8 (b) shows the alphabetic character started by removing the frame shown in drawing 28 (a) by 44g of frame removal sections.
[0083] As explained above, in this example, it has the following descriptions 1-6.
[0084] 1) Even if it is the block frame and free-formatting frame of a tabular format with an inclination, a letter face can be extracted and removed certainly.
[0085] 2) Since mask processing is performed instead of thinning, the processing time can be shortened.
[0086] 3) Since mask processing is performed instead of thinning, the linearity which a subject-copy image has can be maintained and, thereby, straight-line detection can be performed easily.
[0087] 4) Since the ratio to the die length of the length of the horizontal die length divides into two or more parts the part which became beyond a predetermined value in the rectangle part surrounded by the frame in four sides in a free-formatting frame or the block frame of a tabular format, even when there is an inclination, a frame can be extracted and removed certainly.
[0088] 5) Since it dissociates for every range of a single-character frame certainly even if two or more frames are connected through the alphabetic character, since matching processing is performed when retrieval of striping goes wrong, a frame can be extracted certainly and can be removed.
[0089] 6) A frame can be more certainly extracted by considering as the start point of frame retrieval of the thinnest part inside the rectangle segment judged to be a frame in the case of the extract of a frame.
[0090] By the way, these people have proposed the **** image extractor previously shown in drawing 29 by Japanese Patent Application No. No. 103257 [ five to ]. The connection pattern extract section 1 extracts the partial pattern with which the pixel and the pixel are connected among this drawing from the image which consists of straight-line parts, such as a frame or a ruled line, an alphabetic character, a graphic form, or a notation. The straight-line extract section 2 with thick line breadth extracts straight lines, such as a frame with the thick line breadth which is equivalent to the length of the size of a connection pattern, and the horizontal die length for every partial pattern, and a ruled line, by asking for projection of the partial pattern extracted in the connection pattern extract section 1. The straight-line extract section 3 with thin line breadth extracts straight lines which were not extracted in the straight-line extract section 2 with thick line breadth, such as a frame and a ruled line. The frame extract section 4 extracts four sides which constitute a frame from two or more extracted straight lines. That is, the frame extract section 4 detects a frame by asking for the straight line of the maximum outline of a partial pattern among the straight lines extracted in the straight-line extract section 2 with thick line breadth, or the straight-line extract section 3 with thin line breadth. If the connection pattern attribute adjunct 5 investigates whether straight lines, such as a frame and a ruled line, exist to a connection pattern and does not exist in it, it will add the attribute of the pattern (Ha) which consists of an alphabetic character, a graphic form, a notation, or its part. Moreover, the connection pattern attribute adjunct 5 adds the attribute of the alphabetic character in contact with straight lines, such as the attribute of straight-line pattern (b)s, such as a frame and a ruled line, a frame, and a ruled line, a graphic form, a notation, or a

part of [ the ] pattern (b)s after straight-line separation of a frame, a ruled line, etc. according to whether an alphabetic character, a graphic form, a notation, or some of its partial patterns exist.
[0091] The separation section 6 computes the width of face of straight lines, such as a frame and a ruled line, and separates straight lines, such as a frame and a ruled line, from a connection pattern according to it. The intersection calculation section 7 computes the intersection when straight lines, such as a frame and a ruled line, meet with an alphabetic character, a graphic form, or a notation. The intersection matching section 8 matches between intersections according to the conditions of the continuity of the segment based on the distance between the segments which constitute an alphabetic character, a graphic form, or a notation, and a direction etc. The interpolation section 9 interpolates the field of the alphabetic character in straight lines, such as a frame and a ruled line, a graphic form, or a notation by connecting between the intersections matched in the intersection matching section 8. The connective check section 10 checks the connectivity of a pattern, reverse-using the information that the pattern obtained in the interpolation section 9 had connected from the first.
[0092] When connectivity cannot be checked in the connective check section 10, it matches with processing of the intersection matching section 8 by extending the conditions of matching of return and an intersection. And about the matched intersection, it interpolates in the interpolation section 9 and a connective check is again performed in the connective check section 10. The connection pattern integrated section 11 unifies both in consideration of the size when unifying the alphabetic character and graphic form which were interpolated in the interpolation section 9, or the pattern and the above-mentioned pattern of a notation while it unifies the pattern (Ha) which consists of an alphabetic character, a graphic form, a notation, or its part and extracts an alphabetic character, a graphic form, or a notation.
[0093] When this interpolates correctly the part which removed straight lines, such as a frame or a ruled line, certainly, and was removed from the image which consists of straight-line parts, such as a frame or a ruled line, an alphabetic character, a graphic form, or a notation, the alphabetic character, graphic form, or notation of high quality can be extracted. For this reason, in a character reader etc., recognition rates, such as an alphabetic character, a graphic form, and a notation, can be raised remarkably.
[0094] Moreover, the character string extract section 12 extracts the character string which is obtained through the connection pattern integrated section 11 and by which labeling was carried out. The average character-size calculation section 13 computes an average character size based on the extracted character string. The small separation stroke extract section 14 extracts a small separation stroke using an average character size, area conditions, and height conditions. The character-pitch calculation section 15 computes an average character pitch by making distance between circumscription rectangles into a pitch about what was not judged to be a small separation stroke. The average character-size pitch integrated section 16 unifies an alphabetic character based on information, such as an extracted small separation stroke, an average character pitch, an average character size, and a size variance. The reliability integrated section 17 computes the distance of the extracted small separation stroke and the character pattern located in the right and left, quantifies it as reliability of integration of the ratio of those distance, and when reliability is high, it unifies. Moreover, the simple recognition processing section 18 performs processing for a handwriting alphabetic character (figure) with an overhang. That is, to the pattern located in a small separation stroke and its right and left, and the pattern at the time of unifying them further, the simple recognition processing section 18 identifies linear density, an inclination, and a character size simply, and unifies an alphabetic character.
[0095] However, in the image extractor of drawing 29 , since the intersection of an alphabetic character and a frame is computed and matched after removing a frame, when the frame width is quite larger than the character width of the alphabetic character which touches the frame as shown [ especially ] in drawing 30 (a), or when the frame inclines as shown in drawing 31 (a), the information about an alphabetic character will decrease extremely. In the case of drawing 31 (a), as shown in this drawing (b), the information about most alphabetic characters will be lost. Moreover, since the field perpendicularly larger horizontally [ in the case of drawing 31 (a) / as shown in this drawing (b) / the frame extracted ] than original is occupied, the information about most alphabetic characters as well as the case of

drawing 30 will be lost. Hatching shows a frame among drawing 30 and drawing 31 .
[0096] Therefore, only on the conditions used in the intersection matching section 8 of drawing 29 , although "7" does not touch "1" in fact like drawing 33 when the intersection which "1" is divided into right and left like drawing 32 , and is not matched is generated, when a mistake is made in computing the intersection of a frame and an alphabetic character and it matches it, matching which was mistaken like drawing 34 may be performed. Moreover, after separation of a frame, it may not interpolate, even after the character pattern which lapped with the frame like drawing 35 recombining to the pattern with which connectivity was checked, and the conditions of matching are insufficient only by connective check. Among drawing 32 - drawing 35 , (a) shows a subject-copy image and (b) shows the alphabetic character finally started by matching in the intersection matching section 8 of drawing 29 . In addition, a "black dot" shows the intersection of a frame and an alphabetic character among drawing 32 and drawing 33 R> 3. Moreover, drawing 34 shows matching which "9" mistook and drawing 35 shows the case where "8" is not interpolated.
[0097] Next, the 2nd example of the image extractor which becomes this invention is explained with drawing 36 and drawing 37 . In this example, this invention is applied to the processing which starts an alphabetic character from a black-framed document. that is, the case where there are two or more letter faces separated in this example with the rectangle which the size or the location of one letter face do not understand, began to see from contact or its letter face to the letter face, and a handwriting alphabetic character is written -- the pattern of a letter face to an alphabetic character and an alphabetic character part -- a single character -- one character is started.
[0098] In this example The connection pattern extract section 21 which shows an image extractor to drawing 36 , the straight line / frame extract section 22, the connection pattern attribute adjunct 23, the contact frequency calculation section 51, the frame separation section 24, the intersection calculation section 25-1, 25-2, the intersection matching sections 55 and 61 shown in drawing 37 , the within the limit alphabetic character extract section 56, the within the limit alphabetic character interpolation section 32, The character string extract section 12-1, 12-2, the average character-size pitch integrated section 16-1, 16-2, two or more connectivity check section 57-1, 57-2, the connection pattern integrated section 34, the re-interpolation judging section 63, the average character-size pitch integrated section 16-3, the reliability integrated section 17 And it consists of the simple recognition processing section 18. 3 intersection calculation section 25-1 consists of an alphabetic character / frame intersection calculation section 25a, the intersection calculation section 52-2 of a within the limit alphabetic character, the intersection calculation section 53-1 of an outside the limit alphabetic character, and an alphabetic character / frame intersection judging section 54-1. 2 intersection calculation section 25-2 consists of the intersection calculation section 53-2 of an alphabetic character / frame intersection calculation section 25a, and an outside the limit alphabetic character, and an alphabetic character / frame intersection judging section 54-2.
[0099] In drawing 36 , the connection pattern extract section 21 has the labeling section which extracts the pattern connected in 8 connection by labeling from the input configuration signal with which it pretreated. The pattern of "8 Connection" means the pattern connected in length, width, or eight slanting directions. General technique can be used for the above-mentioned labeling. In this example, the partial pattern obtained in the labeling section is either of the alphabetic characters in contact with a part of alphabetic character which does not touch the frame with which the alphabetic character does not touch, and a frame, alphabetic character, or a frame. A frame is extracted in order to observe only the alphabetic character which distinguishes these partial patterns and touches the frame. Moreover, since it is needed by the processing which the size of the partial pattern obtained by labeling mentions later, the coordinate of the angle of the rectangle obtained by approximating a partial pattern to a rectangle is computed during processing of labeling.
[0100] A straight line / frame extract section 22 extracts a straight line / frame with thick line breadth, and a straight line/frame with thin line breadth. As shown in drawing 38 , a straight line / frame extract section 22 has projection section 22a, straight-line detecting-element 22b, four-side detecting-element 22c, 22d of trace sections, four-side detecting-element 22e, and 22f of frame extract sections.

[0101] Projection section 22a projects the partial pattern extracted by the connection pattern extract section 21 on horizontal and a perpendicular direction. Straight-line detecting-element 22b detects a horizontal line and a vertical line from the horizontal projection obtained by projection section 22a, and vertical projection. Four-side detecting-element 22c detects four sides of the rectangle which consists of a straight line detected by straight-line detecting-element 22b. 22d of trace sections pursues a line by the n run length technique mentioned later in order to ask for a straight line with the thin line breadth which was not able to be detected by straight-line detecting-element 22b and four-side detecting-element 22c. Four-side detecting-element 22e detects four sides from the rectangle which consists of a straight line with the thin line breadth for which it asked by 22d of trace sections. 22f of frame extract sections extracts a frame from the rectangle for which it asked by the four-side detecting elements 22c and 22e.
[0102] The connection pattern attribute adjunct 23 has the attribute adjunct which adds the attribute of a "frame", "a character pattern or its part", and "a frame, a character pattern or some of its contact patterns" about the connection pattern extracted by the connection pattern extract section 21 based on the extract result of a straight line / frame extract section 22, as shown in drawing 36 .
[0103] The contact frequency calculation section 51 computes contact frequency from the degree of contact of a frame, the width of character and an alphabetic character, and a frame. A character width is larger than frame width, and contact frequency is small when the alphabetic character and the frame do not touch in comparison. on the other hand, a character width and frame width -- abbreviation -- it is the same, or contact frequency is large when there is comparatively much contact in an alphabetic character and a frame. When contact frequency is small, after separating a frame in the frame separation section 24, the intersection of an alphabetic character and a frame is computed in 2 intersection calculation section 25-2. Moreover, when contact frequency is large, the intersection of an alphabetic character and a frame is computed in 3 intersection calculation section 25-1.
[0104] In addition, the frame separation section 24 separates a frame from the connection pattern with which the attribute was given as a contact pattern of a frame, or an alphabetic character and a frame by the connection pattern attribute adjunct 23. As shown in drawing 39 , the frame separation section 24 performs labeling again about the pattern from which **** calculation section 24a which computes the width of face of the side for a frame part, frame removal section 24b which removes a frame, and a frame were removed, and has frame noise rejection section 24c and 24d of attribute adjuncts which remove a pattern with a small area as a noise. 24d of attribute adjuncts adds the attribute of a contact character pattern to the pattern which remains even if it removes a frame to the pattern with which an attribute was not given in the connection pattern attribute adjunct 23, and if a frame is removed, they will add the attribute of only a frame to the pattern which does not remain at all.
[0105] When contact frequency is large, a frame is not separated but the alphabetic character / frame intersection calculation section 25a in 3 intersection calculation section 25-1 compute the intersection of a frame and an alphabetic character about a contact character pattern. The intersection calculation section 52-1 of a within the limit alphabetic character asks for the area of the alphabetic character segment at the same time it computes an intersection within the limit by searching for an alphabetic character segment in the within the limit direction from the location about the intersection of all alphabetic characters and frames. However, when an alphabetic character segment breaks off during retrieval of an alphabetic character segment, it judges with the computed intersection not being an intersection of an alphabetic character and a frame. Next, the intersection calculation section 53-1 of an outside the limit alphabetic character asks for the area of the alphabetic character segment at the same time it computes an intersection outside the limit by searching for an alphabetic character segment to some extent with frame width in the outside the limit direction from the location about the intersection of all alphabetic characters and frames. With the outside the limit direction, it is the direction where the within the limit direction is opposite, the direction which goes outside from a frame is pointed out, and an intersection distant from a frame is pointed out with an intersection outside the limit. An alphabetic character / frame intersection judging section 54-1 judges with the intersection not being an intersection of an alphabetic character and a frame while it will consider that an alphabetic character segment is a noise and it will remove it, if the sum of the area of the alphabetic character segment searched for in the

intersection calculation section 52-1 and the intersection calculation section 52-2 outside the limit within the limit is below a predetermined threshold.
[0106] On the other hand, when contact frequency is small, after separating a frame in the frame separation section 24, an intersection is computed in 2 intersection calculation section 25-2. The alphabetic character / frame intersection calculation section 25a in 2 intersection calculation section 25-2 compute the intersection of a frame and an alphabetic character about a contact character pattern. The intersection calculation section 53-2 of an outside the limit alphabetic character asks for the area of the alphabetic character segment at the same time it computes an intersection outside the limit by searching for an alphabetic character segment to some extent with frame width in the outside the limit direction from the location about the intersection of all alphabetic characters and frames. An alphabetic character / frame intersection judging section 54-2 judges with the intersection not being an intersection of an alphabetic character and a frame while it will consider that an alphabetic character segment is a noise and it will remove it, if the area of the alphabetic character segment searched for in the intersection calculation section 52-2 outside the limit is below a predetermined threshold.
[0107] When contact frequency is large and calculation of an intersection is performed in 3 intersection calculation section 25-1, an intersection is matched in 3 intersection matching section 55 shown in drawing 37 . On the other hand, when contact frequency is small and calculation of an intersection is performed in 2 intersection calculation section 25-2, an intersection is matched in 2 intersection matching section 61 shown in drawing 37 .
[0108] 3 intersection matching section 55 searches for the directivity of the alphabetic character segment based on an intersection distant from the frame obtained in 3 intersection calculation section 25-1, the intersection of an alphabetic character and a frame, and three kinds of intersections within the limit. Moreover, the intersection of an alphabetic character and a frame is matched according to the conditions of the continuity of the alphabetic character segment based on the distance and the direction of [ between this directivity and the alphabetic character segment in an intersection within the limit ]. The within the limit alphabetic character extract section 56 extracts from a frame the alphabetic character segment which constitutes between the intersections of a corresponding alphabetic character and a corresponding frame. The character string extract section 12-1 extracts the character string which is obtained through the within the limit alphabetic character extract section 56 and by which labeling was carried out. The average character-size pitch integrated section 16-1 unifies an alphabetic character provisionally based on information, such as an extracted small separation stroke, an average character pitch, an average character size, and a size variance. When two or more connectivity check section 57-1 checks connectivity based on the number of connection segments, the number of holes, the Euler number, etc. and connectivity is checked, processing progresses to the connection pattern integrated section 34. However, interpolation processing which returns to processing of 3 intersection matching section 55, performs matching during an intersection which recovers connectivity, connects the character pattern separated by having connected the intersection further matched in the within the limit alphabetic character interpolation section 32, and having removed the letter face when connectivity cannot be checked, and interpolates an alphabetic character field within the limit is performed. After finishing processing in this within the limit alphabetic character interpolation section 32, processing progresses to the connection pattern integrated section 34.
[0109] 2 intersection matching section 61 searches for the directivity of the alphabetic character segment based on the intersection of an alphabetic character and a frame and two kinds of intersections within the limit which were obtained in 2 intersection calculation section 25-2. Moreover, the intersection of an alphabetic character and a frame is matched according to the conditions of the continuity of the alphabetic character segment based on the distance and the direction of [ between this directivity and the alphabetic character segment in an intersection within the limit ]. The within the limit alphabetic character interpolation section 32 performs interpolation processing which connects the character pattern separated by having connected the matched intersection and having removed the letter face. The character string extract section 12-2 extracts the character string which is obtained through the within the limit alphabetic character interpolation section 32 and by which labeling was carried out. The

average character-size pitch integrated section 16-2 unifies an alphabetic character provisionally based on information, such as an extracted small separation stroke, an average character pitch, an average character size, and a size variance. When two or more connectivity check section 57-2 checks connectivity based on the number of connection segments, the number of holes, the Euler number, etc. and connectivity is checked, processing progresses to the connection pattern integrated section 34. However, interpolation processing which returns to processing of 2 intersection matching section 61, performs matching during an intersection which recovers connectivity, connects the character pattern separated by having connected the intersection further matched in the within the limit alphabetic character interpolation section 32, and having removed the letter face when connectivity cannot be checked, and interpolates an alphabetic character field within the limit is performed. After finishing processing in this within the limit alphabetic character interpolation section 32, processing progresses to the connection pattern integrated section 34.

[0110] The re-interpolation judging section 63 judges the existence of the re-interpolated character pattern, and when there is nothing, it progresses to processing of the reliability integrated section 17 as it is. On the other hand, when there is a re-interpolated character pattern, after computing an average character size and a pitch finally in the average character-size pitch integrated section 16-3 and unifying an alphabetic character, it progresses to processing of the reliability integrated section 17. The reliability integrated section 17 computes the distance of the extracted small separation stroke and the character pattern located in the right and left, quantifies it as reliability of integration of the ratio of those distance, and when reliability is high, it unifies. Moreover, the simple recognition processing section 18 performs processing for a handwriting alphabetic character (figure) with an overhang. That is, to the pattern located in a small separation stroke and its right and left, and the pattern at the time of unifying them further, the simple recognition processing section 18 identifies linear density, an inclination, and a character size simply, and unifies an alphabetic character.

[0111] The within the limit alphabetic character interpolation section 32 connects the intersection matched in the intersection matching section 55 or 61, and connects the character pattern separated by having removed the letter face. The within the limit alphabetic character interpolation section 32 has simple interpolation section 32a, crossing calculation section 32b, crossing within the limit interpolation section 32c, 32d of crossing outside the limit interpolation sections, and linear interpolation section 32e, as shown in drawing 40 . Simple interpolation section 32a is interpolated by connecting between the intersections of a one to one correspondence. Crossing calculation section 32b computes the crossing of the direction vector of the profile of the alphabetic character segment in an intersection about the intersection and with eclipse ****** corresponding to re-corresponding to one to many. Crossing within the limit interpolation section 32c connects between intersections, when the crossing for which it asked is in the line breadth of a frame. 32d of crossing outside the limit interpolation sections connects between intersections, when the crossing for which it asked is outside the line breadth of a frame. Linear interpolation section 32e interpolates an alphabetic character segment in a straight line along with a frame, when a crossing cannot be computed.

[0112] The connection pattern integrated section 34 unifies the character pattern to which the attribute of a character pattern was added in the connection pattern attribute adjunct 23 of the character pattern or drawing 36 by which connectivity was checked. As shown in drawing 41 , the connection pattern integrated section 34 has connection pattern temporary integrated section 34a, size check section 34b, and connection pattern integrated section 34c. Connection pattern temporary integrated section 34a performs temporary integration of a connection pattern. Size check section 34b checks the size of the character pattern by which temporary integration was carried out. Connection pattern integrated section 34c unifies a connection pattern further, when the size checked in size check section 34b is not suitable.

[0113] Next, logging processing of the alphabetic character in this example is explained.

[0114] (1) The input configuration to which pretreatment of amendment of an extreme inclination and rotation, removal of a noise, stopgap of "graze", etc. was performed is inputted into the connection pattern extract section 21 of extract drawing 36 of a connection pattern. The labeling section in the connection pattern extract section 21 extracts the pattern connected in length, width, or the eight

directions direction of slant by labeling in order to choose the candidate of a contact alphabetic character regardless of the location of a frame about this input configuration. the partial pattern obtained by such labeling -- (**) -- a frame when the alphabetic character does not touch, and (**) -- it is either of the alphabetic characters in contact with a part of alphabetic character which does not touch a frame, alphabetic character, or (Ha) a frame.

[0115] In addition, since the size of the partial pattern obtained by labeling is needed by next processing like ****, the coordinate of the angle of the rectangle obtained by approximating a partial pattern to a rectangle during processing of labeling is searched for.

[0116] (2) the extract of a frame -- in order to observe only the alphabetic character which touches the frame among the connection patterns extracted as mentioned above, extract a frame in the frame extract section 22. In case a frame is extracted, generally the frequency of an extract is high, and since it is necessary to extract at a high speed, when a straight line/frame with thick line breadth are extracted first and this is not able to extract, a straight line/frame with thin line breadth are extracted next.

[0117] (2a) The size of the extract frame of a straight line/frame with thick line breadth is strange, and even when the alphabetic character touches the frame, it is necessary to extract a frame to stability. Then, projection section 22a of the frame extract section 22 shown in drawing 38 takes projection for every partial pattern obtained by labeling. And in straight-line detecting-element 22b, a ratio with the size in every direction obtained by approximating a projection value and a partial pattern to a rectangle is calculated, and if the ratio is more than a predetermined threshold, it will be judged that it is a long straight line.

[0118] In addition, since this straight line has not only a frame but the case of a stroke of an alphabetic character, it asks for the straight line equivalent to four sides which are satisfied with the straight line of the maximum outline of rectangle-likeness as much as possible.

[0119] Drawing 42 is drawing showing the horizontal projection of a partial pattern and the vertical projection which were obtained by labeling in the connection pattern extract section 21. The partial pattern 41 obtained by labeling has die-length Lx in a longitudinal direction among this drawing, and has die length Ly in a lengthwise direction. Moreover, a partial pattern has vertical Projection VP and the vertical horizontal projection HP.

[0120] Here, it is [ size / of the expedient top of explanation, and an image ] Ph about f (i, j) and horizontal projection of the i-th line in m train xn line and the concentration value of a coordinate (i, j). It is Pv about projection of the perpendicular direction of (i) and eye j train. It is Ph if (j). (i) and Pv (j) is expressed with the following (2) types and (3) types, respectively.

[0121]

[Equation 1]

$$Ph\ (i) = \sum_{j=1}^{m} f\ (i,\ j) \qquad (2)$$

$$Pv\ (j) = \sum_{i=1}^{n} f\ (i,\ j) \qquad (3)$$

[0122] Straight-line detecting-element 22b about the horizontal projection for which it asked by projection section 22a, and vertical projection The die length Lx and Ly of the rectangle of the rectangle coordinate of the partial pattern for which it asked in the connection pattern extract section 21 in every direction is used. They are the vertical die length Lx and the horizontal projection Ph by the following (4) types. They are the horizontal die length Ly and the vertical projection Pv by the ratio with (i), and the following (5) types. It asks for a ratio with (j) and is a threshold THL about these ratios. It compares.

[0123]

[Ph (i)/Lx] >=THL (4)

[Pv (j)/Ly] >=THL (5)

And the ratio in (4) types and (5) types is a threshold THL. If it is above, a partial pattern shall be the

candidate of the straight line which constitutes a frame. That is, as shown in drawing 42 , when the extracted partial pattern is a rectangle, it is the level projection value Ph of the straight-line part. (i) and perpendicular projection value Pv Since (j) becomes the largest and a ratio with the die length Lx and Ly in every direction also becomes large, a straight-line part can be distinguished by (4) types and (5) types.

[0124] Drawing 43 and drawing 44 are drawings which explain more one example of processing of straight-line detecting-element 22b to a detail. Drawing 43 shows the component of a frame and the frame is constituted by a cope box, a drag flask, a left frame, and the right frame. A cope box consists of segments (straight line) i1 and i11. A drag flask consists of segment (straight line) i2 and i22. A left frame consists of segments (straight line) j1 and j11. The right frame consists of segments (straight line) j2 and j22. If the die length of the rectangle of the partial pattern obtained by labeling in every direction is set to Lx and Ly, respectively, the ratio of projection and die length is the predetermined threshold THL. If it is above, it will consider as the candidate of the segment (straight line) which constitutes a frame.
[0125] Drawing 44 shows a flow chart in case software realizes processing of straight-line detecting-element 22b. Steps S1-S7 detect the candidate of a cope box among horizontal lines among this drawing. Steps S11-S17 detect the candidate of a drag flask among horizontal lines. Steps S21-S27 detect the candidate of a left frame among vertical lines. Moreover, steps S31-S37 detect the candidate of the right frame among vertical lines. Here, only actuation of the expedient top of explanation and steps S1-S7 is explained, and explanation of other steps is omitted.
[0126] In drawing 44 , step S1 is set as i= 0. step S2 -- [Ph (i)/Lx] <THL it is -- a ****** is judged. After incrementing i at step S3 as the judgment result of step S2 is NO, it returns to step S2. On the other hand, step S4 sets it as i1=i that the judgment result of SUTTEPPU S2 is YES. next, the step S5 -- [Ph (i)/Lx] >=THL it is -- a ****** is judged. After incrementing i at step S6 as the judgment result of step S5 is YES, it returns to step S5. On the other hand, the candidate of a cope box is obtained by i11=i-1 by setup and these actuation at step S7 as the judgment result of SUTTEPPU S5 is NO.
[0127] The candidate of a drag flask, a left frame, and the right frame is similarly obtained by steps S11-S17, steps S21-S27, and steps S31-S37, respectively. In addition, steps S13 and S16 carry out the decrement of the i, for example.
[0128] Out of the candidate of the horizontal line i detected by straight-line detecting-element 22b, and the candidate of a vertical line j, paying attention to the horizontal line candidate i1 of the maximum outline, i2, and the candidates j1 and j2 of a vertical line, four-side detecting-element 22c calculates the following (6) types and (7) types, and compares with threshold THL', respectively.
[0129]
| i1-i2|/Ly>=THL' (6)
| j1-j2|/Lx>=THL' (7)
Drawing 45 is drawing which explains more one example of processing of four-side detecting-element 22c to a detail, and shows a flow chart in case software realizes processing. Step S41 judges whether the above-mentioned (6) formula is materialized among this drawing. When the judgment result of step S41 is YES, step S42 judges whether the above-mentioned (7) formula is materialized. Processing progresses that step S41 or the judgment result of S42 is NO to processing of 22d of trace sections. On the other hand, when the judgment result of step S42 is YES, it considers that step S43 is a straight-line part from which a partial pattern constitutes a frame, and a straight-line part is detected.
[0130] 22f of frame extract sections extracts a frame based on the detection result of four-side detecting-element 22c. That is, if a frame will be extracted based on this if a straight-line part is detected by four-side detecting-element 22c, and a straight-line part is not detected, a frame will be extracted by repeating the above-mentioned processing paying attention to other candidates. If a candidate satisfies the above-mentioned (6) formula and (7) types, it will be regarded as the straight line which constitutes a frame, and if not satisfied, specifically paying attention to other candidates, the above-mentioned processing is repeated.
[0131] Paying attention to the skeleton order, it calculates whether the candidate i of a how many

horizontal line or the candidate j of a vertical line exists continuously from a skeleton, and let the value be the criteria of the line breadth of each side, after asking for the straight line which constitutes the frame of a frame as mentioned above.
[0132] Drawing 46 (a) is drawing showing an example of a frame with the thick line breadth extracted as mentioned above. A skeleton 152 is obtained to the extracted frame 151 among this drawing, and the line breadth of each side is 2 pixels in this example.
[0133] (2b) Extract a straight line/frame with thin line breadth paying attention to the partial pattern which was not able to be computed in the extract processing of a straight line/frame with the thick line breadth of drawing 46 (a) in which the straight line/frame with thin line breadth carried out the extract above.
[0134] Drawing 46 (b) is drawing showing an example of a frame with thin line breadth, and shows a frame 153 and its skeleton 154. As shown in this drawing, the pattern which line breadth is about 1 pixel minute, and the irregularity by an inclination etc. has produced is contained in a frame with thin line breadth. In order to extract to stability a frame with the thin line breadth shown in this drawing, this example is searched for a frame as follows.
[0135] That is, as shown in drawing 47 , even if the irregularity by the inclination has arisen, the straight-line length who names the "n line run length" which can detect a straight line is defined as the extract of a frame.
[0136] Although it will be divided into a short straight line with the usual run length in the long straight line which irregularity as shown in drawing 46 (b) has produced since the number of pixels which is continuing horizontally or vertically is calculated, the run length connected by 8 connection among certain n lines is calculated as a n line run length in the n line run length shown in drawing 47 . The value of n is determined in the magnitude of an inclination, and the more an inclination is large, the more it enlarges n. In the case of n= 1, it is equivalent to the usual run length.
[0137] Drawing 47 can extract a straight line as a horizontal line for 7 pixels, although the case of n= 3 was shown and irregularity has arisen in this case.
[0138] 22d of trace sections of the frame extract section 22 makes a start point the location which took down the perpendicular from two or more points, such as both ends of the straight line which constitutes a rectangle from a rectangle coordinate of the partial pattern obtained in the labeling section of the connection pattern extract section 21, and the middle point, to the interior of a frame, and contacted the partial pattern, and they search it along with a partial pattern with the above-mentioned n line run length in right and left or the vertical direction from a start point.
[0139] Stability can be asked for a straight line with thin line breadth even if the alphabetic character has overflowed the frame by using such technique.
[0140] Drawing 48 and drawing 49 are drawings which explain more one example of processing of 22d of trace sections to a detail, and show a flow chart in case software realizes processing. Drawing 48 R> 8 shows the processing which asks for a horizontal n line run length, and drawing 49 shows the processing which asks for a vertical n line run length. In drawing 48 and drawing 49 , the die length of Nhi and a vertical line shall be expressed [ a retrieval start point ] with Nvj for the die length of (ii, jj), and a horizontal line.
[0141] Step S51 is set as Nhi=0, i=ii, and j=jj among drawing 48 . It judges whether step S52 has a label in f (i+1, j). After incrementing Nhi and i at step S53 as the judgment result of step S52 is YES, processing is returned to step S52. On the other hand, if the judgment result of step S52 is NO, it will judge whether it is <=(j+1) jj+n/2 at step S54. It judges whether a label is that the judgment result of step S54 is YES in f (i+1, j+1) at step S55. If the judgment result of step S55 is YES, after incrementing j at step S56, processing will be returned to step S53. When step S54 or the judgment result of S55 is NO, it judges whether it is >=(j-1) jj-n/2 at step S57. It judges whether a label is that the judgment result of step S57 is YES in f (i+1, j-1) at step S58. If the judgment result of step S58 is YES, after carrying out the decrement of the j at step S59, processing will be returned to step S53. When step S57 or the judgment result of S58 is NO, the processing which asks for a horizontal n line run length is ended.
[0142] Step S61 is set as Nvj=0, i=ii, and j=jj among drawing 49 . It judges whether step S62 has a label

in f (i, j+1). After incrementing Nvj and j at step S63 as the judgment result of step S62 is YES, processing is returned to step S62. On the other hand, if the judgment result of step S62 is NO, it will judge whether it is <=(i+1) ii+n/2 at step S64. It judges whether a label is that the judgment result of step S64 is YES in f (i+1, j+1) at step S65. If the judgment result of step S65 is YES, after incrementing i at step S66, processing will be returned to step S63. When step S64 or the judgment result of S65 is NO, it judges whether it is >=(i-1) ii-n/2 at step S67. It judges whether a label is that the judgment result of step S67 is YES in f (i-1, j+1) at step S68. If the judgment result of step S68 is YES, after carrying out the decrement of the i at step S69, processing will be returned to step S63. When step S67 or the judgment result of S68 is NO, the processing which asks for a vertical n line run length is ended.

[0143] In the candidate i of a horizontal line, the candidate j of Nhi and a vertical line sets to Nvj the die length of the straight line from which four-side detecting-element 22e was obtained as mentioned above. Moreover, using the die length Lx and Ly of the rectangle of the rectangle coordinate of the partial pattern for which it asked in the labeling section of the connection pattern extract section 21 in every direction, it asks for the ratio of the horizontal die length Ly and the candidate Nvj of a vertical line with the following (8) types by the ratio of the vertical die length Lx and the candidate Nhi of a horizontal line, and the following (9) types, and is a threshold THL about these ratios. It compares.

[0144]

[Nhi/Lx] >=THL (8)

[Nvj/Ly] >=THL (9)

And the ratio in (8) types and (9) types is a threshold THL. If it is above, he shall be the candidate of the straight line which constitutes a frame for each candidate.

[0145] Drawing 50 is drawing which explains more one example of processing of four-side detecting-element 22e to a detail, and shows a flow chart in case software realizes processing. The candidate j of Nhi and a vertical line of the candidate i of the inside of this drawing and a horizontal line shall be Nvj. Moreover, die length with a rectangle in every direction than the rectangle coordinate of the partial pattern obtained on the occasion of labeling is set to Lx and Ly, respectively. this drawing (a) -- setting -- step S71 -- Nhi/Lx>=THL it is -- a ****** is judged. If the judgment result of step S71 is YES, it will be judged that step S72 is suitable as a candidate of the straight line from which this candidate i constitutes a frame. On the other hand, if the judgment result of step S71 is NO, processing will be started to the candidate of the next horizontal line. moreover, this drawing (b) -- setting -- step S75 -- Nvj/Ly>=THL it is -- a ****** is judged. If the judgment result of step S75 is YES, it will be judged that step S76 is suitable as a candidate of the straight line from which this candidate j constitutes a frame. On the other hand, if the judgment result of step S75 is NO, processing will be started to the candidate of the following vertical line.

[0146] Out of the candidate of the horizontal line i detected by straight-line detecting-element 22b, and the candidate of a vertical line j, paying attention to the horizontal line candidate i1 of the maximum outline, i2, and the candidates j1 and j2 of a vertical line, four-side detecting-element 22c calculates the following (10) types and (11) types, and compares with threshold THL', respectively.

[0147]

| i1-i2|/Ly>=THL' (10)

| j1-j2|/Lx>=THL' (11)

22f of frame extract sections extracts a frame like the case where it describes above, based on the detection result of four-side detecting-element 22e. That is, if a frame will be extracted based on this if a straight-line part is detected by four-side detecting-element 22e, and a straight-line part is not detected, a frame will be extracted by repeating the above-mentioned processing paying attention to other candidates. If a candidate satisfies the above-mentioned (10) formula and (11) types, it will be regarded as the straight line which constitutes a frame, and if not satisfied, specifically paying attention to other candidates, the above-mentioned processing is repeated. In addition, the reference value of line breadth is calculated as a difference of the max obtained during the frame extract, and the minimum position coordinate.

[0148] Thus, after asking for one letter face, it is also possible to scan the extracted letter face over all

images, and to newly extract a letter face, after taking matching with the partial pattern extracted by the connection pattern extract section 21.
[0149] Moreover, the size ratio of the rectangle which approximated the partial pattern for which it asked in the connection pattern extract section 21, and the extracted letter face is computed, and after the computed size ratio narrows down only to the partial pattern in a predetermined threshold and taking matching with a letter face [ finishing / an extract ], it is also possible to newly extract a letter face.
[0150] Furthermore, when judged with it being the pattern which consists only of an alphabetic character in an attribute addition means by which extract the partial pattern which exists within the limits of right and left or the upper and lower sides by the size of the extracted letter face, and all the extracted partial patterns explain it below, a means to rejudge the letter face by which the extract was carried out [ above-mentioned ] to be a character pattern may be established. When such a means is established, for example, even if it is an alphabetic character with the rectangle part which is equivalent to a frame like the kanji "country", it can prevent judging a part of alphabetic character as a frame accidentally.
[0151] (3) the result of a frame extract [ in / in the attribute addition connection pattern attribute adjunct 23 to a connection pattern / the frame extract section 22 ] -- being based -- each connection pattern -- (**) of a degree - (Ha) -- divide like and add the attribute of a frame and an alphabetic character to a frame pattern, a character pattern, and the partial pattern of an alphabetic character, respectively.
(b) contact pattern (**) of an alphabetic character and a frame -- the partial pattern of a frame pattern (Ha) alphabetic character and an alphabetic character, in addition the pattern which was not able to extract a frame add the attribute of a "character pattern" or "a part of character pattern." Moreover, the pattern to which the attribute was added as mentioned above is the connection pattern integrated section 34 mentioned later, and it is decided whether start by the pattern independent, or regard as a part of character pattern, and to unify with other patterns.
[0152] (4) Among separation drawing 39 of a frame, **** calculation section 24a of the frame separation section 24 calculates the line breadth reference value +2 (it is plus 1 to plus 1 and an inner profile in an outside profile) acquired in consideration of the irregularity of a profile, and the irregularity of an inner profile outside the frame at the time of a frame extract as line breadth, and determines the coordinate value of a profile and an inner profile outside a frame from the max under the skeleton computed in the frame extract, or extract / the minimum location.
[0153] Drawing 51 is drawing showing the skeleton of a frame with thick line breadth, and a frame with thin line breadth, an outside profile, an inner profile, and line breadth, in "a thick line", an outside profile and a "dotted line" show an inner profile, and "a thin line" shows a skeleton. This drawing (a) shows a frame with thick line breadth, and this drawing (b) shows a frame with thin line breadth. **** calculation section 24a determines the coordinate value of a profile and an inner profile outside a frame about a frame with thick line breadth, and a frame with thin line breadth, as shown in this drawing.
[0154] If **** calculation section 24a shall show system of coordinates to drawing 52 , width of face will set that the coordinate of a profile is the frame coordinate of a cope box w1 in (i1, i11) outside a cope box, and it will set (i1-1, i11+1) and neighboring width of face to w1+2. Moreover, width of face sets that the coordinate of a profile is the frame coordinate of a drag flask w2 in (i2, i22) outside a drag flask, and (i2 +1, i22-1) and neighboring width of face are set to w2+2. Similarly width of face sets that the coordinate of a profile is the frame coordinate of a left frame w3 in (j1, j11) outside a left frame, and (j1-1, j11+1) and neighboring width of face are set to w3+2. Furthermore, width of face sets that the coordinate of a profile is the frame coordinate of the right frame w4 in (j2, j22) outside the right frame, and (j2+1, j22-1) and neighboring width of face are set to w4+2.
[0155] If it returns to explanation of the frame separation section 24 of drawing 39 , frame removal section 24b will remove the pattern which exists between an outside profile and an inner profile with the coordinate value of a profile and an inner profile outside the frame for which it asked in **** calculation section 24a.
[0156] Frame noise rejection section 24c performs labeling again to the pattern from which the frame was removed, and removes the pattern which remains as some frames according to the descriptions, like

area is small for every label. If they add the attribute of the pattern in which it is shown that that in which the pattern exists is a part of contact character pattern or alphabetic character and a frame is removed even if 24d of attribute adjuncts removes a frame by the connection pattern attribute adjunct 23 paying attention to the pattern with which an attribute was not added, they will add the attribute of the pattern in which it is shown that that whose pattern is lost is only a frame.
[0157] In addition, it is judged whether it is unified with other contact character patterns or a part of alphabetic character in the connection pattern integrated section 34 or after a part of contact character pattern or alphabetic character interpolates an alphabetic character part within the limit so that it may mention later, it is used as an independent pattern.
[0158] (5) When smaller than the threshold which computes average frame width, the inclination of a frame, and the rate of a contact alphabetic character to whole sentence number of letters, and has either among the computed values, the contact frequency calculation section 51 shown in calculation drawing 36 of contact frequency judges with contact frequency being small, and when larger than a certain threshold, it judges with contact frequency being large.
[0159] Drawing 53 is drawing which explains more one example of processing of the contact frequency calculation section 51 to a detail, and shows a flow chart in case software realizes processing. Step S501 receives the attribute information from the continuation pattern attribute adjunct 23 etc., step S502 computes the average frame width Wave based on this, and step S503 computes inclination muki of a frame. Step S504 will judge whether it is Wave>Th.Wave, if Th.Wave is made into the threshold of average frame width. Step S505 computes the whole sentence number of letters znum as the decision result of step S504 is YES. Step S506 will judge whether it is muki>Th.muki, if Th.muki is made into the threshold of the inclination of a frame. Step S507 computes the number snum of contact alphabetic characters as the decision result of step S506 is YES. Step S508 computes rate sbaz=snum/znum to the whole sentence number of letters of the number of contact alphabetic characters. Step S509 will judge whether it is sbaz>Th.sbaz, if Th.sbaz is comparatively made into the threshold of sbaz. The decision result of step S509 is judged to be YES with the contact alphabetic character frequency SH being large at step S510. On the other hand, steps S504 and S506 or the decision result of S509 is judged to be NO with the contact alphabetic character frequency SH being small at step S511.
[0160] (6) Compute the intersection of the pattern and frame for a part of contact character pattern or alphabetic character in the alphabetic character / frame intersection calculation section 25a of calculation drawing 36 of an alphabetic character / frame intersection. That is, using the outside profile and inner profile for which it asked in the frame separation section 24, an outside straight line and an inner profile are asked for an inside straight line by 1 pixel by 1 pixel at an outside profile, the point of changing from white or white to black is searched for from black in that a part of the straight line and contact character pattern, or alphabetic character crosses, and this is extracted as an intersection.
[0161] Drawing 54 is drawing which explains more one example of processing of an alphabetic character / frame intersection calculation section 25a to a detail, and shows a flow chart in case software realizes processing. In this case, the coordinate inside is1 and an inner profile is set to iu1 for the coordinate of the outside of a profile outside a cope box using the same system of coordinates as drawing 52 . The following processings are performed setting the coordinate inside is2 and an inner profile to iu1 for the coordinate of the outside of a profile outside a drag flask, setting the coordinate inside js1 and an inner profile to ju1 for the coordinate of the outside of a profile outside a left frame, and using the coordinate inside js2 and an inner profile as ju1 for the coordinate of the outside of a profile outside the right frame.
[0162] Drawing 54 shows the processing which computes a profile and the intersection of an alphabetic character outside a cope box. Among this drawing, step S81 sets an intersection number K as K= 0, and step S82 sets j as js1 <=j<=js2. That is, the following steps S83-S89 are performed about j=js1-js2 by setup of step S82. It judges whether step S83 has a label in f (is1, j). After making j increase that there is no label in f (is1, j), and the judgment result of step S83 is YES at step S84, processing is returned to step S83. On the other hand, if the judgment result of step S83 is NO, the coordinate on the left-hand side of an intersection will be set as j at step S85. Moreover, it judges whether step S86 has a label in f

(is1, j). After making j increase that the judgment result of step S86 is NO at step S87, processing is returned to step S86. On the other hand, if the judgment result of step S86 is YES, the coordinate on the right-hand side of an intersection will be set as j at step S88. Then, K is made to increase at step S89. After following j=js1-js2 in steps S83-S89, the processing which computes a profile and the intersection of an alphabetic character outside a cope box is ended.

[0163] In addition, since the same processing as drawing 54 can perform calculation of the intersection of a profile, an inner profile, and an alphabetic character outside calculation of the inner profile of a cope box, and the intersection of an alphabetic character, a drag flask, a left frame, and the right frame, the illustration and explanation are omitted on these specifications.

[0164] (7) The intersection calculation section 52-1 of the within the limit alphabetic character in the intersection calculation 3 intersection calculation section 25-1 of a within the limit alphabetic character asks for the area of an alphabetic character segment, pursuing the profile of an alphabetic character segment in the within the limit direction from the intersection of a frame and an alphabetic character. However, when an alphabetic character segment stops existing while pursuing to frame width, it judges with the intersection obtained not being an intersection of an alphabetic character and a frame.

[0165] Drawing 55 is drawing which explains more one example of processing of the intersection calculation section 52-1 of a within the limit alphabetic character to a detail, and shows a flow chart in case software realizes processing. Step S521 receives the intersection from an alphabetic character / frame intersection calculation section 25a etc., and initial setting is performed at step S522. At step S522, the intersection of ww=0, an alphabetic character, and a frame is set as A0 and (B0) character-width mw=B0-A0 and area Mtikai=mw of an alphabetic character segment. If step S523 is made into ww=ww +1 and step S524 sets frame width to w, it will judge whether it is ww<w. Processing progresses that the decision result of step S524 is NO to processing of the intersection calculation section 53-1 of an outside the limit alphabetic character at step S525. On the other hand, when the decision result of step S524 is YES, the intersection of ww point is pursued at step S526. Moreover, step S527 judges whether an intersection (Aww, Bww) within the limit exists.

[0166] Drawing 56 shows the example of a trace of an intersection within the limit, and a trace is performed as the drawing Nakaya mark shows. Among drawing 56 , the intersection of a frame and an alphabetic character and a "black square" express an intersection within the limit, and, as for hatching, a "black dot" expresses a black pixel. When the decision result of step S527 is NO, it judges with (A0, B0) not being the intersections of an alphabetic character and a frame at step S528, and processing progresses to step S525. That is, as shown in drawing 57 , when an intersection does not exist within the limit, it judges with the obtained intersection (A0, B0) not being an intersection of an alphabetic character and a frame. On the other hand, when the decision result of step S527 is YES, it judges whether it is mw**Bww-Www at step S529. If the decision result of step S529 is NO, processing will progress to step S525. However, when the decision result of step S529 is YES, processing is returned to step S523 as mw=Bww-Www and Mtikai=Mtikai+mw at step S530.

[0167] (8) The intersection calculation section 53-1 of the outside the limit alphabetic character in the intersection calculation 3 intersection calculation section 25-1 of an outside the limit alphabetic character asks for the area of an alphabetic character segment, pursuing the profile of an alphabetic character segment in the outside the limit direction from the intersection of a frame and an alphabetic character. However, when an alphabetic character segment stops existing while pursuing to frame width, it judges that the intersection of the point is an intersection distant from a frame.

[0168] Drawing 58 is drawing which explains more one example of processing of the intersection calculation section 53-1 of an outside the limit alphabetic character to a detail, and shows a flow chart in case software realizes processing. Step S 541-1 receives an intersection etc. through the intersection calculation section 53-1 of a within the limit alphabetic character, and initial setting is performed at step S542. At step S542, the intersection of ww=0, an alphabetic character, and a frame is set as A0 and (B0) character-width mw=B0-A0 and area Mtooi=mw of an alphabetic character segment. If step S543 is made into ww=ww +1 and step S544 sets frame width to w, it will judge whether it is ww<w. Processing progresses that the decision result of step S544 is NO to processing of an alphabetic

character / frame intersection judging section 54-1 at step S545. On the other hand, when the decision result of step S544 is YES, the intersection of ww point is pursued at step S546. Moreover, step S547 judges whether an intersection (Aww, Bww) outside the limit exists.
[0169] Drawing 59 shows the example of a trace of an intersection outside the limit, i.e., an intersection distant from a frame, and a trace is performed as the drawing Nakaya mark shows. Among drawing 59 , the intersection of a frame and an alphabetic character and a "black triangle" express an intersection outside the limit, and, as for hatching, a "black dot" expresses a black pixel. When the decision result of step S547 is NO, processing progresses to step S545. However, as shown in drawing 60 , it is in the middle of a trace, and when an alphabetic character segment stops existing, it judges with it being the intersection where the intersection (A0, B0) of the point is distant from a frame. On the other hand, when the decision result of step S547 is YES, it judges whether it is mw**Bww-Www at step S548. If the decision result of step S548 is NO, processing will progress to step S545. However, when the decision result of step S548 is YES, processing is returned to step S543 as mw=Bww-Www and Mtooi=Mtooi+mw at step S549.
[0170] In addition, in processing of the intersection calculation section 53-2 of the outside the limit alphabetic character in 2 intersection calculation section 25-2, except that step S541-2 are performed instead of step S541-1 in drawing 58 , it is the same as processing of the intersection calculation section 53-1 of an outside the limit alphabetic character. Moreover, step S541-2 receive an intersection etc. through the alphabetic character / frame intersection calculation section 25a in not the intersection calculation section 53-1 of a within the limit alphabetic character but 2 intersection calculation section 25-2.
[0171] (9) The alphabetic character / frame intersection judging section 54-2 in the alphabetic character / frame intersection judging section 54-1 in the judgment 3 intersection calculation section 25-1 of an alphabetic character / frame intersection, or 2 intersection calculation section 25-2 judge with the intersection obtained as it is below a threshold with the sum of the area of an alphabetic character segment within the limit and outside the limit not being an intersection of an alphabetic character and a frame.
[0172] Drawing 61 is drawing which explains more one example of processing of an alphabetic character / frame intersection judging section 54-1, or the alphabetic character / frame intersection judging section 54-2 to a detail, and shows a flow chart in case software realizes processing. Step S551 receives an intersection etc. from the intersection calculation section 53-1 of an outside the limit alphabetic character, or 53-2, and step S552 will set it as Mwa=Mtikai+Mtooi, if the sum of the area of an alphabetic character segment within the limit and outside the limit is set to Mwa. Step S553 will judge whether it is Mwa>Th.menseki, if the threshold of the sum of area is set to Th.menseki. It is judged with an intersection (A0, B0) not being an intersection of a frame and an alphabetic character at step S554 as the decision result of step S553 is NO, and processing progresses to step S556. On the other hand, it is judged with an intersection (A0, B0) being an intersection of a frame and an alphabetic character at step S555 as the decision result of step S553 is YES, and processing progresses to step S556. Step S556 supplies the information about the obtained intersection to an alphabetic character / frame intersection calculation section 25a.
[0173] (10) between the candidates of the intersection which 3 intersection matching section 55 shown in matching drawing 37 of three intersections matches -- the directivity of an intersection distant from a frame, an intersection within the limit, and the intersection of an alphabetic character and a frame to a pattern -- computing -- abbreviation -- if the same direction is computed, match these intersections. Even if it is the case where directivity is not in agreement, when satisfying the distance during an intersection within the limit, and the conditions of continuity correspondence, the intersection of the alphabetic character and frame is matched. Since the condition judging of matching is performed to an intersection within the limit, right matching can be performed.
[0174] Drawing 62 is drawing showing matching of an intersection in case directivity is in agreement, among this drawing, the intersection of a frame and an alphabetic character and a "black triangle" express an intersection distant from a frame, and, as for a black pixel and a "black dot", hatching

expresses directivity, as for an arrow head. On the other hand, drawing 63 is drawing showing matching of an intersection in case directivity is not in agreement, among this drawing, the intersection of a frame and an alphabetic character and a "black triangle" express an intersection distant from a frame, and, as for a black pixel and a "black dot", hatching expresses directivity, as for an arrow head.

[0175] Drawing 64 is drawing which explains more one example of processing of 3 intersection matching section 55 to a detail, and shows a flow chart in case software realizes processing. Step S561 receives the information about the intersection from 3 intersection calculation section 25-1 etc., and step S562 sets frame width as w. Step S563 searches for the directivity H1 of a pattern from the intersection 1 (A, B) and the intersection 1 (An, Bn) within the limit of an alphabetic character and a frame, and the intersection 1 (At, Bt) distant from a frame. Step S564 searches for the directivity H2 of a pattern from the intersection 2 (C, D) and the intersection 2 (Cn, Dn) within the limit of an alphabetic character and a frame, and the intersection 2 (Ct, Dt) distant from a frame. It judges whether step S565 is H1**H2. Therefore, correspondence of directivity is checked at steps S563-S565.

[0176] Step S566 asks that the decision result of step S565 is NO for the middle point T1 of the intersection 1 within the limit from T1= (An+Bn) / 2. Moreover, step S567 asks for the middle point T2 of the intersection 2 within the limit from T2= (Cn+Dn) / 2. Step S568 finds the distance KN during an intersection within the limit from KN=ABS (T1-T2). It judges whether step S569 is KN<w. Step S570 finds the distance L1 and L2 during an intersection from L1= (Cn-Bn) and L2= (An-Dn) as the decision result of step S569 is NO. Step S It judges whether 571 areL1>0, and judges whether it is L1<w/2 at step S572 as the decision result of step S571 is YES. When the decision result of step S571 or step S572 is NO, it judges whether it is L2>0 at step S573. It judges whether it is L2<w/2 at step S574 as the decision result of step S573 is YES. Therefore, distance correspondence is checked at steps S566-S574.

[0177] When the decision result of step S573 or step S574 is NO, it judges whether it is An<Cn<Bn at step S575. When the decision result of step S575 is NO, it judges whether it is An<Dn<Bn at step S576. Processing returns that the decision result of step S576 is NO to step S563. Therefore, continuity correspondence is checked at step S575 and step S576.

[0178] In addition, when steps S565, S569, S572, and S574 or the decision result of S576 is YES, processing is put into processing of the within the limit alphabetic character extract section 56 at step S577.

[0179] Since the condition judging of matching is performed to an intersection within the limit, right matching can be performed as shown in drawing 65 and drawing 66 . Drawing 65 shows the case where matching is performed by the intersection within the limit, a black pixel and a "black dot" show the intersection of a frame and an alphabetic character, and, as for a "black square", hatching shows an intersection within the limit. It is shown among drawing 65 that an actual pattern is shown, (b) is matched on the intersection of an alphabetic character and a frame, and there is (no a) **, and (c) shows that matching is performed by the intersection within the limit. Moreover, drawing 66 shows the case where matching is not performed by the intersection within the limit, a black pixel and a "black dot" show the intersection of a frame and an alphabetic character, and, as for a "black square", hatching shows an intersection within the limit. Among drawing 66 , (a) shows an actual pattern and (b) shows that matching is not performed by the intersection within the limit. (11) between the candidates of the intersection which 2 intersection matching section 61 shown in matching drawing 37 of two intersections matches -- the directivity of an intersection distant from a frame, an intersection within the limit, and the intersection of an alphabetic character and a frame to a pattern -- computing -- abbreviation -- if the same direction is computed, match these intersections. The intersection is matched when satisfying the distance during the intersection of an alphabetic character and a frame, and the conditions of continuity correspondence, even if it is the case where directivity is not in agreement.

[0180] Drawing 67 is drawing which explains more one example of processing of 2 intersection matching section 61 to a detail, and shows a flow chart in case software realizes processing. Step S581 receives the information about the intersection from 2 intersection calculation section 25-2 etc., and step S582 sets frame width as w. Step S583 searches for the directivity H1 of a pattern from the intersection 1 (A, B) of an alphabetic character and a frame, and the intersection 1 (At, Bt) distant from a frame. Step

S584 searches for the directivity H2 of a pattern from the intersection 2 (C, D) of an alphabetic character and a frame, and the intersection 2 (Ct, Dt) distant from a frame. It judges whether step S585 is H1**H2. Therefore, correspondence of directivity is checked at steps S583-S585.

[0181] Step S586 asks that the decision result of step S585 is NO for the middle point T1 of the intersection 1 of an alphabetic character and a frame from T1= (A+B) / 2. Moreover, step S587 asks for the middle point T2 of the intersection 2 of an alphabetic character and a frame from T2= (C+D) / 2. Step S588 finds the distance K during an intersection from KN=ABS (T1-T2). It judges whether step S589 is K<w. Step S590 finds the distance L1 and L2 during an intersection from L1= (C-B) and L2= (A-D) as the decision result of step S589 is NO. Step S It judges whether 591 areL1>0, and judges whether it is L1<w/2 at step S592 as the decision result of step S591 is YES. When the decision result of step S591 or step S592 is NO, it judges whether it is L2>0 at step S593. It judges whether it is L2<w/2 at step S594 as the decision result of step S593 is YES. Therefore, distance correspondence is checked at steps S586-S594.

[0182] When the decision result of step S593 or step S594 is NO, it judges whether it is A<C<B at step S595. When the decision result of step S595 is NO, it judges whether it is A<D<B at step S596. Processing returns that the decision result of step S596 is NO to step S583. Therefore, continuity correspondence is checked at step S595 and step S596.

[0183] In addition, when steps S585, S589, S592, and S594 or the decision result of S596 is YES, processing is put into processing of the within the limit alphabetic character interpolation section 32 at step S597.

[0184] (12) The within the limit alphabetic character extract section 56 shown in extract drawing 37 of a within the limit alphabetic character judges the matched intersection and the range surrounded within the limit to be an alphabetic character component. What was judged to be an alphabetic character component is extracted, and except it, since it is a frame, it is removed. That is, in the case of the pattern shown in drawing 68 (a), the alphabetic character component shown in (b) is extracted. In addition. The alphabetic character component from which the black pixel was extracted for hatching among drawing 68 , and the intersection of an alphabetic character and a frame and crepe were extracted for the black dot is shown.

[0185] Drawing 69 is drawing which explains more one example of processing of the within the limit alphabetic character extract section 56 to a detail, and shows a flow chart in case software realizes processing. Step S601 receives the information about the intersection matched from 3 intersection matching section 55, and step S602 computes the straight lines L1 and L2 passing through during the matched intersection 1 (A1, B1) and an intersection 2 (A2, B-2). Step S603 computes the straight line L3 of a frame with an intersection 1, and step S604 computes the straight line L4 of a frame with an intersection 2. It judges whether step S605 has some which are surrounded by the black pixel within the limit in four straight lines L1, L2, L3, and L4. When the decision result of step S605 is YES, an alphabetic character segment is extracted at step S606, and it moves from processing to processing of the character string extract section 12-1. On the other hand, when the decision result of step S605 is NO, a frame is removed at step S607 and processing progresses to step S608.

[0186] (13) In the within the limit alphabetic character interpolation section 32 shown in interpolation drawing 37 and drawing 40 of a within the limit alphabetic character, divide, when the intersection supports one to one, and when one to many is supported, connect smoothly during a contact character pattern or some of its intersections, and interpolate the alphabetic character part which was missing by having separated the frame.

[0187] Simple interpolation section 32a shown in drawing 40 is a means to connect smoothly during the contact character pattern matched with one to one, or some of its intersections.

* NOTICES *

**JPO and NCIPI are not responsible for any damages caused by the use of this translation.**

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

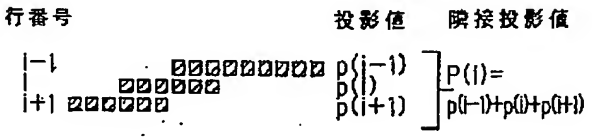
---

DRAWINGS

---

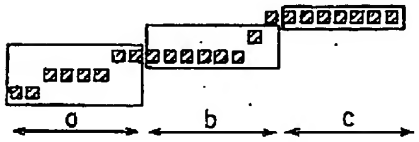
[Drawing 3]

隣接投影法を説明する図



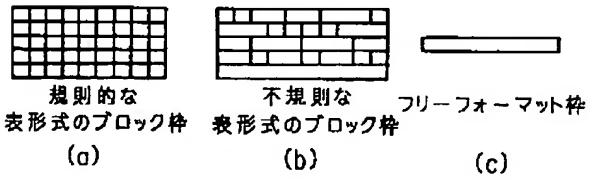

[Drawing 5]

矩形線分の検出を説明する図



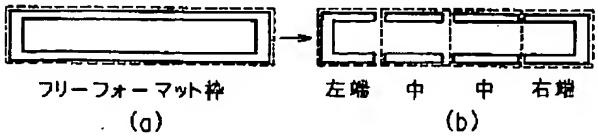

[Drawing 18]

処理の対象となる枠の種類を示す図



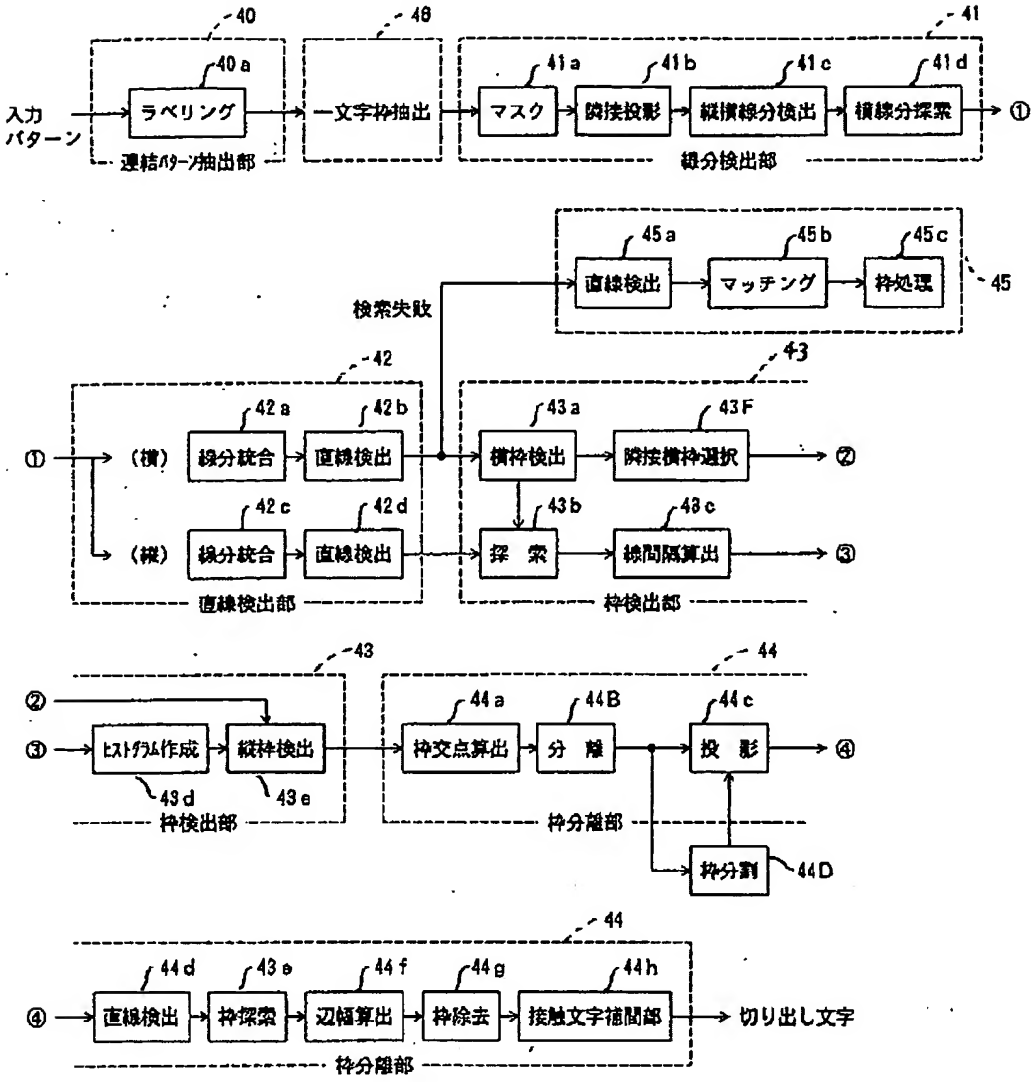

[Drawing 19]

フリーフォーマット枠の分割を説明する図



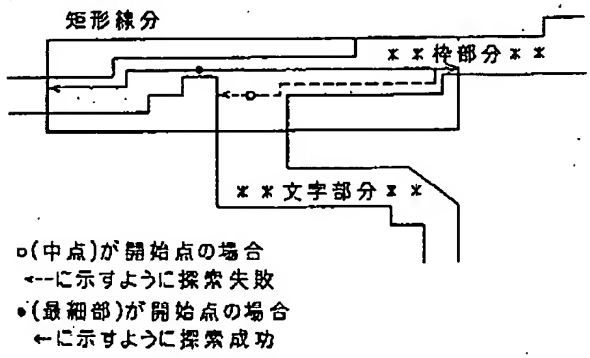

[Drawing 1]

本発明になる画像抽出装置の第１実施例を示すブロック図



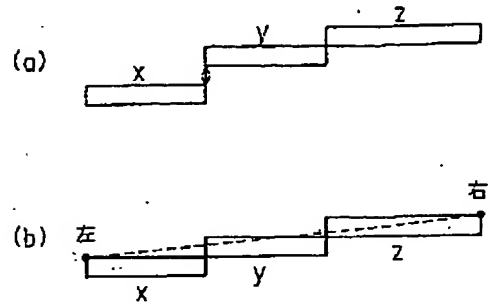

[Drawing 7]

探索の開始点を説明する図



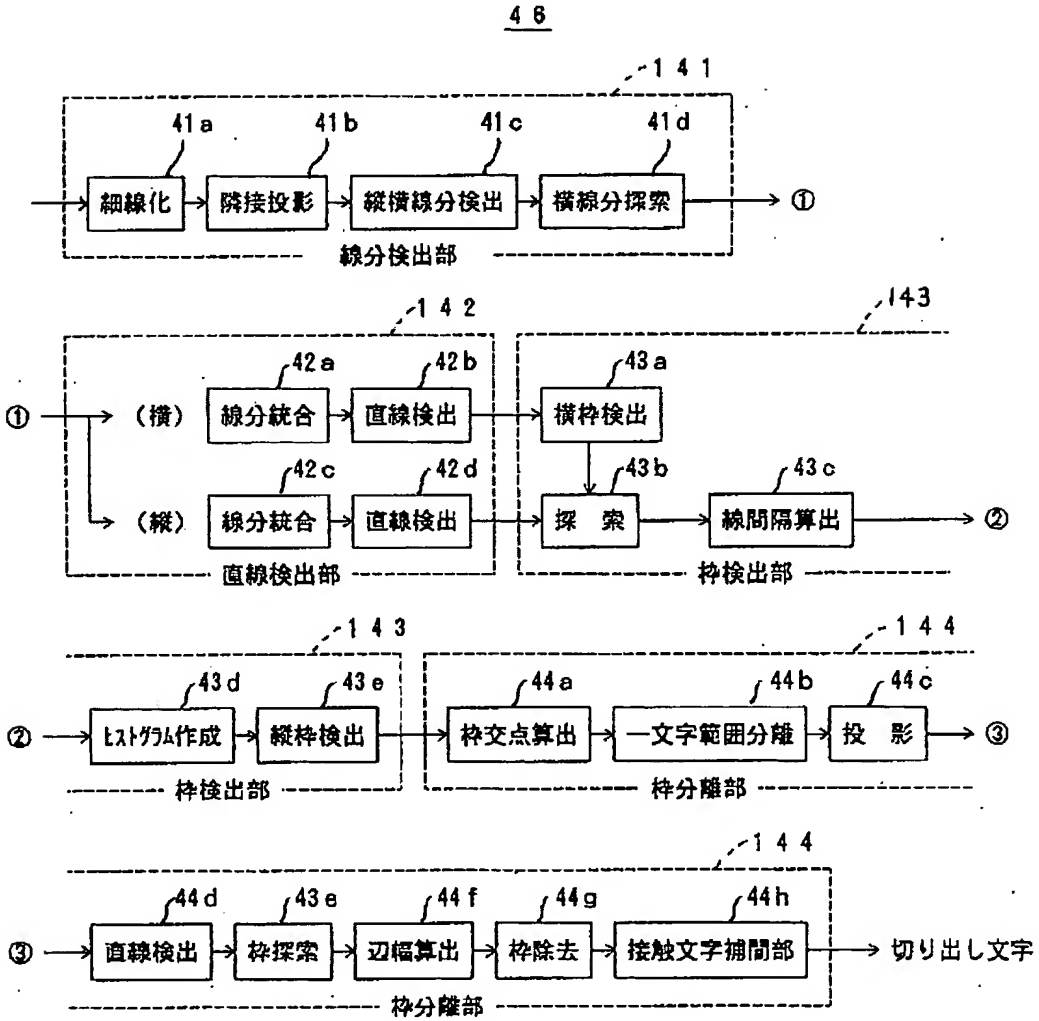

[Drawing 9]

矩形線分の連結及び直線の傾きを説明する図



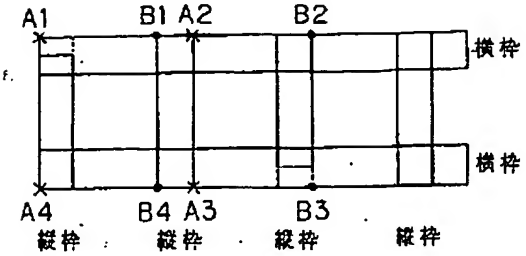

[Drawing 2]

一文字枠抽出部の一実施例を示すブロック図

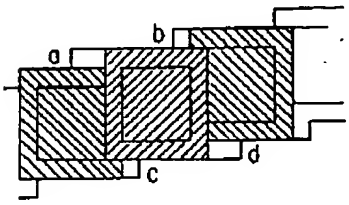



[Drawing 12]

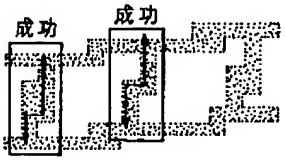
一文字枠への分離を説明する図




[Drawing 13]
枠の除去範囲を説明する図

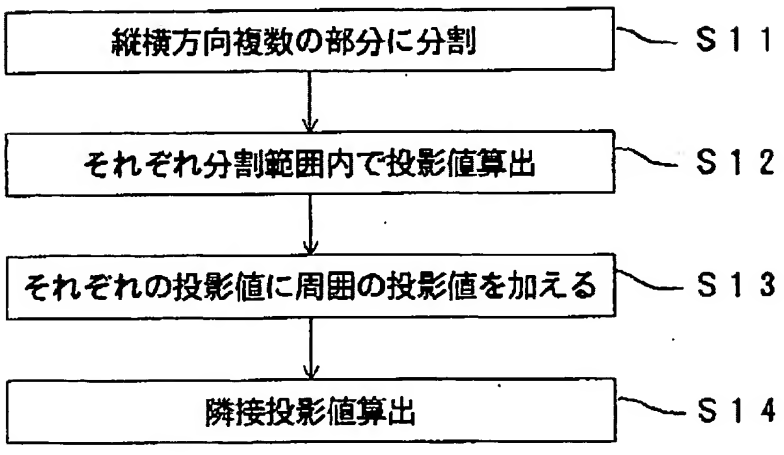



[Drawing 26]
線分の探索を説明する図



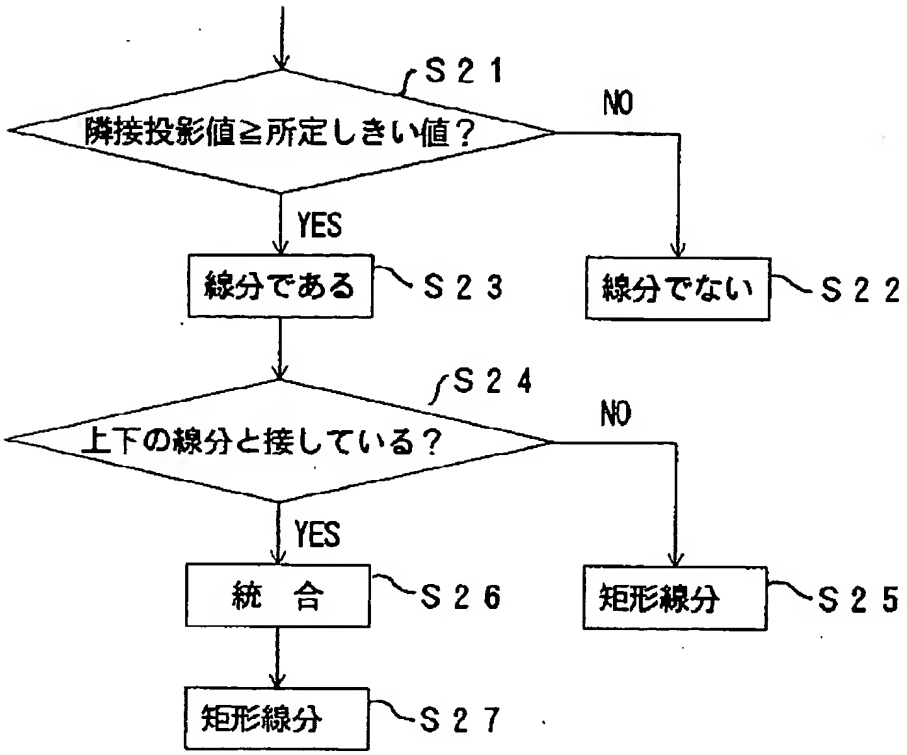

[Drawing 4]

隣接投影部４１ｂの処理を説明するフローチャート



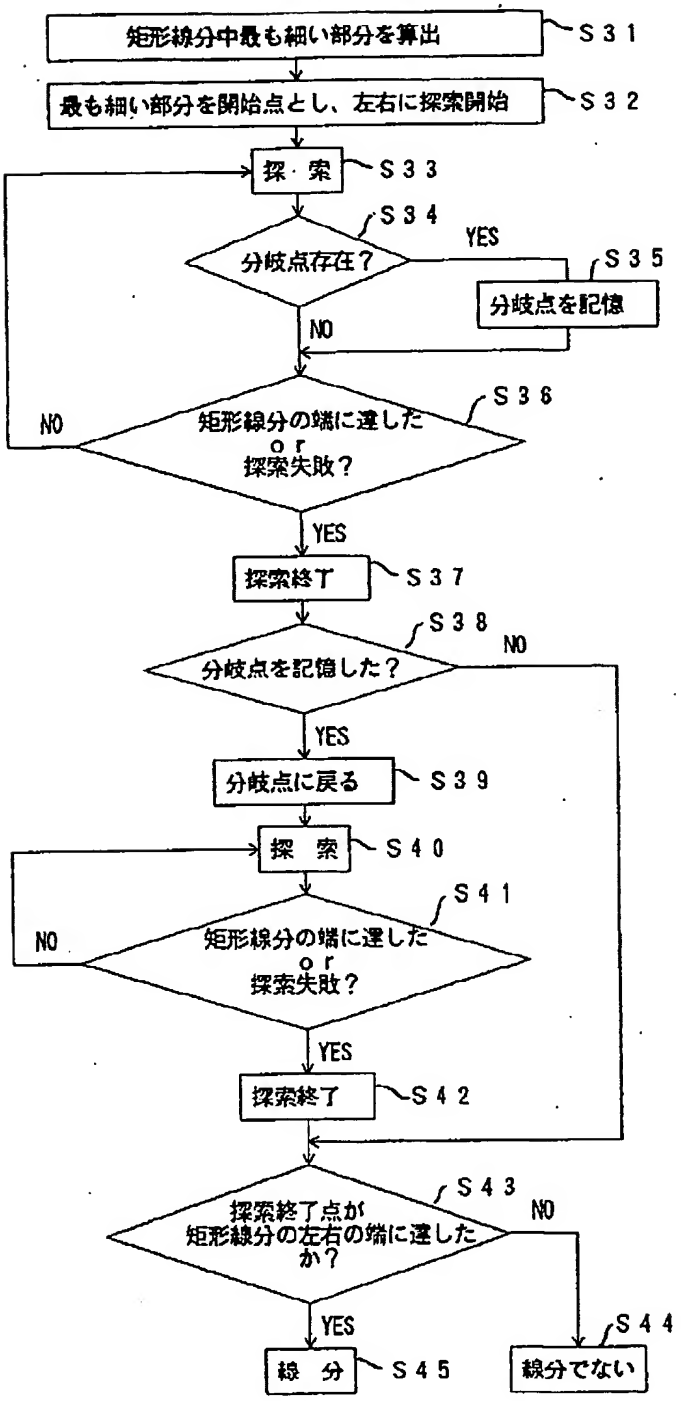

[Drawing 6]

縦横線分検出部４１ｃの処理を説明するフローチャート



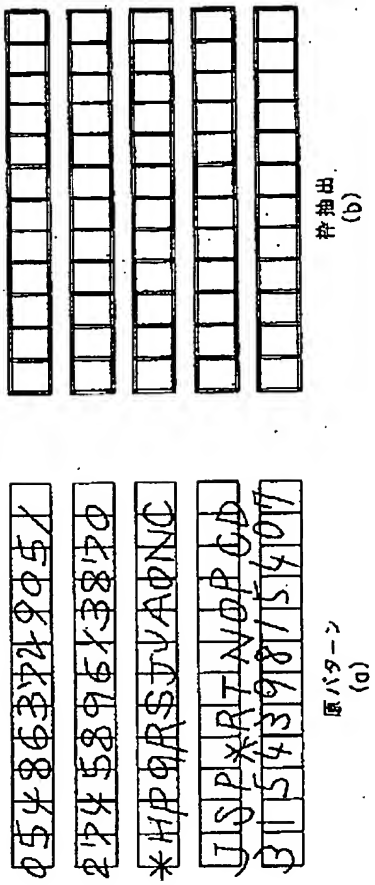

[Drawing 8]

横線分選択部４１ｄの処理を説明するフローチャート



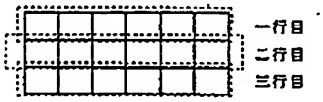

[Drawing 14]

枠の抽出を説明する図



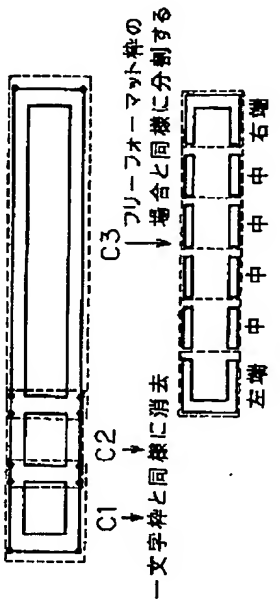

[Drawing 20]

変形式のブロック枠におけるブロック枠の重なりを説明する図

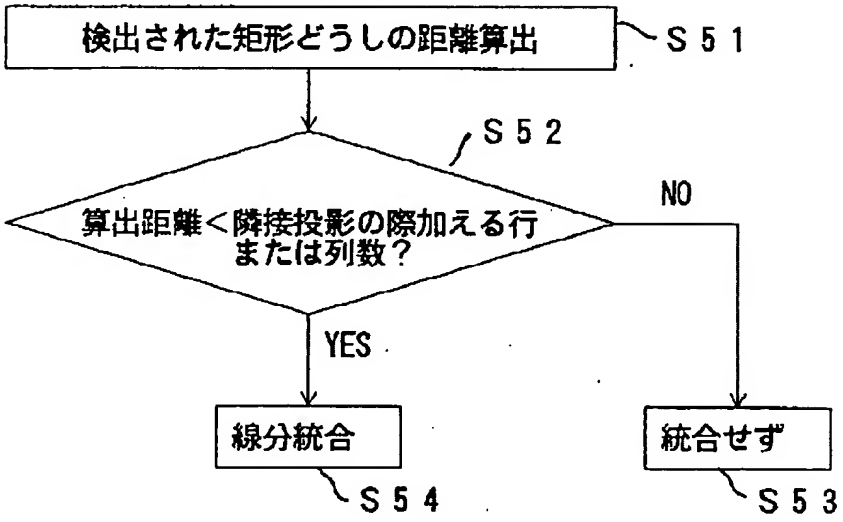



[Drawing 22]

ブロック枠の四辺を枠で囲まれた矩形部分への
分離を説明する図

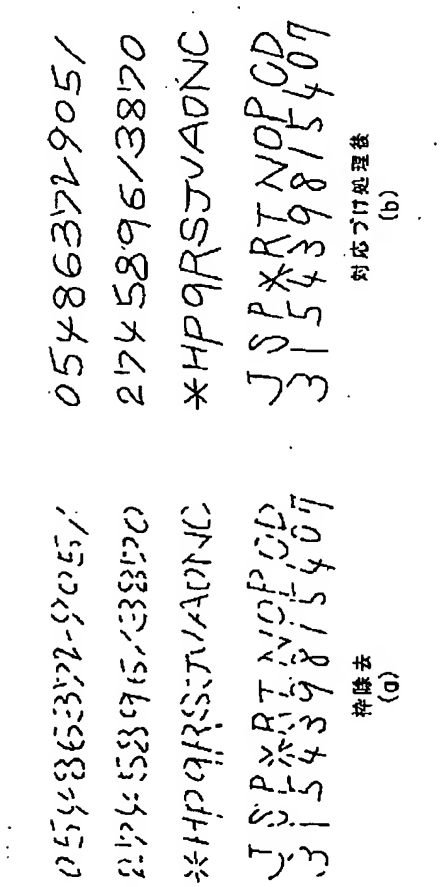



[Drawing 10]
線分統合部４２ａ，４２ｃの処理を説明するフローチャート



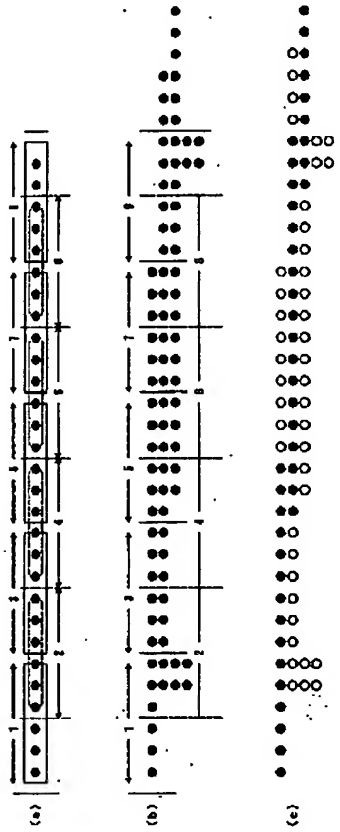

[Drawing 15]

文字の切り出し及び補間を説明する図



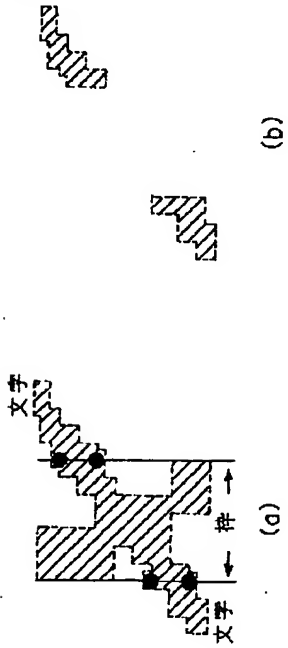

[Drawing 16]

マスクにより構成分を抽出する場合を説明する図



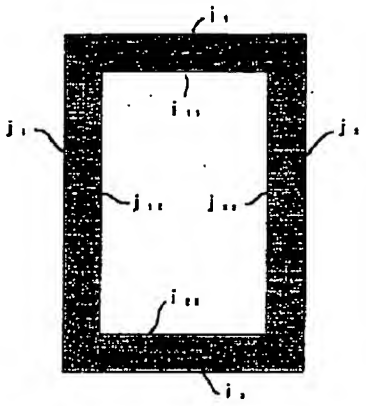
[Drawing 32]

1が左右に分離して対応付けができない例を説明する図



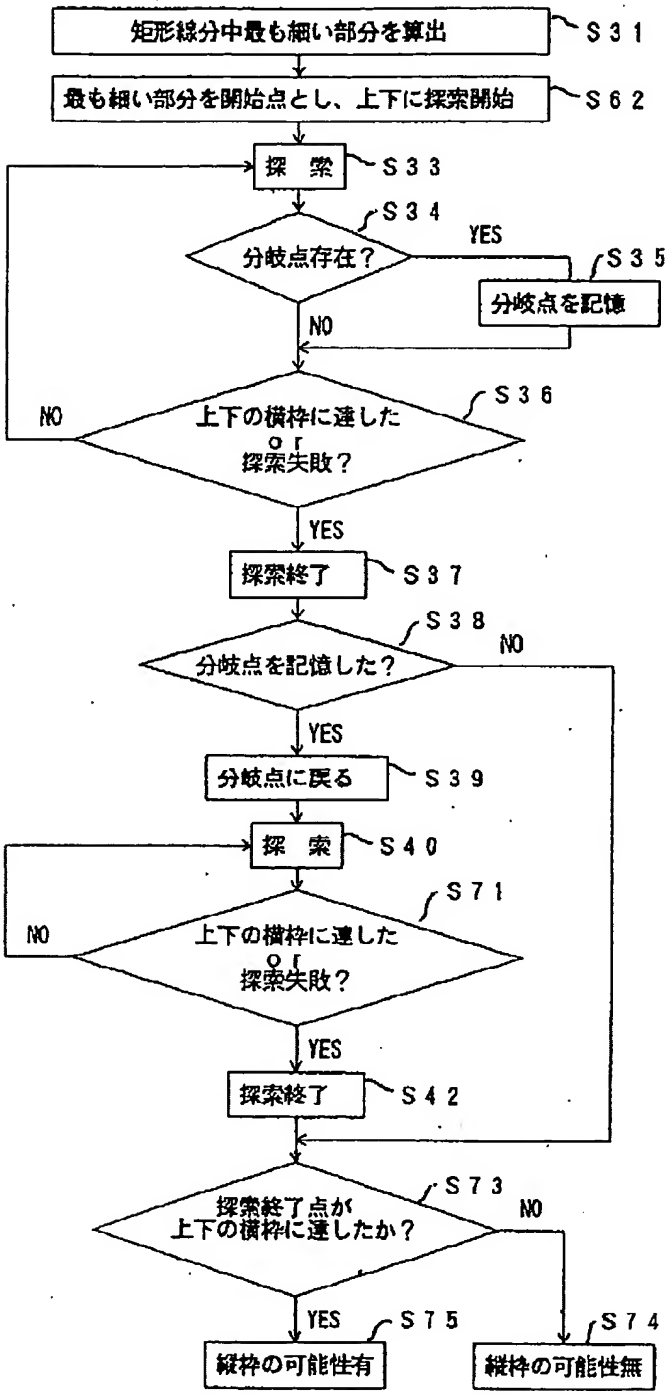

[Drawing 43]

枠の構成要素を示す図



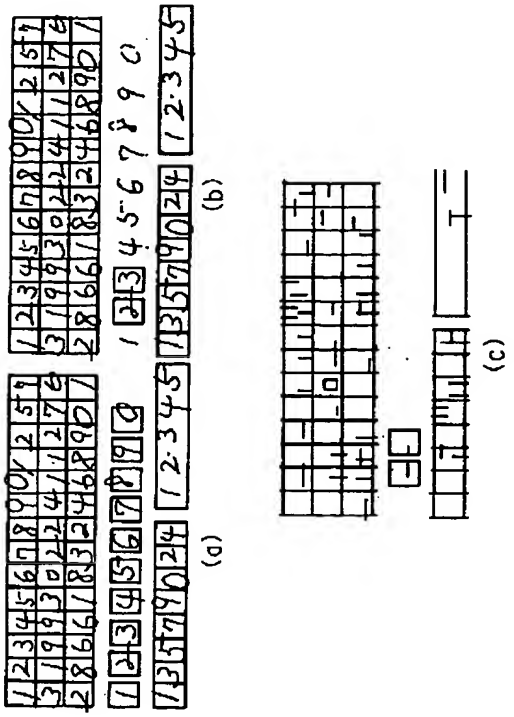
[Drawing 11]

探索部４３ｂの処理を説明するフローチャート



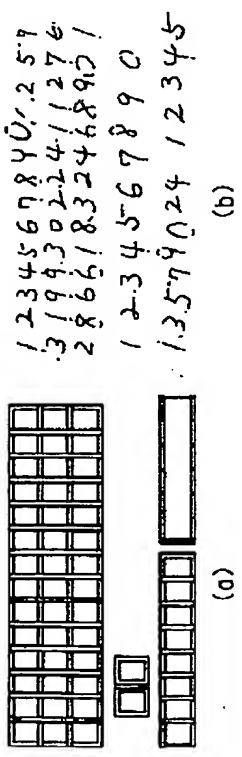

[Drawing 27]

第1実施例により文字を切出した場合の一例を示す図



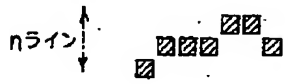
[Drawing 28]

第1実施例により文字を切出した場合の一例を示す図



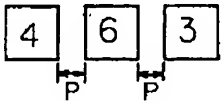
[Drawing 47]

第2実施例における枠探索の手法を示す図



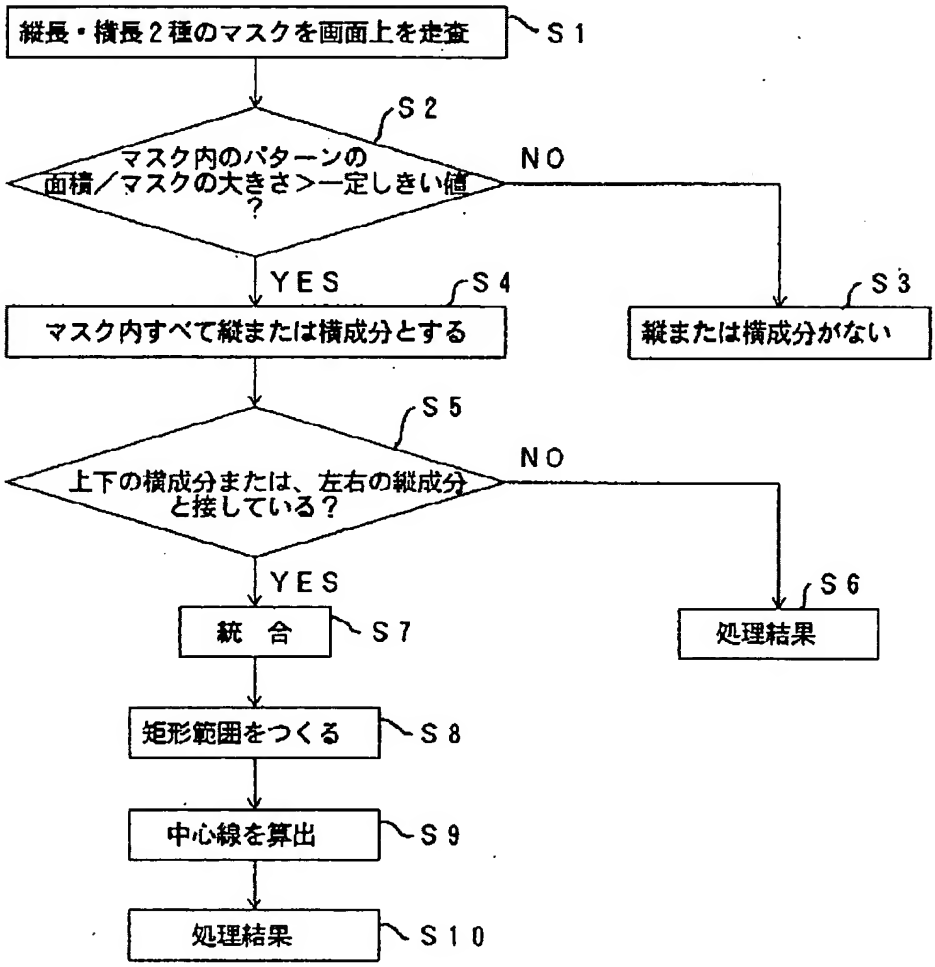

[Drawing 81]

ピッチの算出方法を説明する図



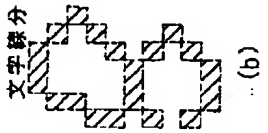

[Drawing 17]

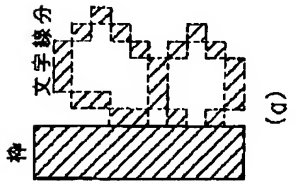
マスク処理部４１Ａの処理を説明するフローチャート



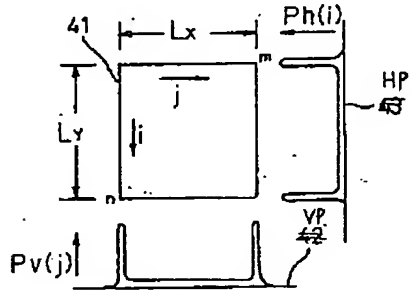

[Drawing 35]

8が再補間されない例を説明ずる図



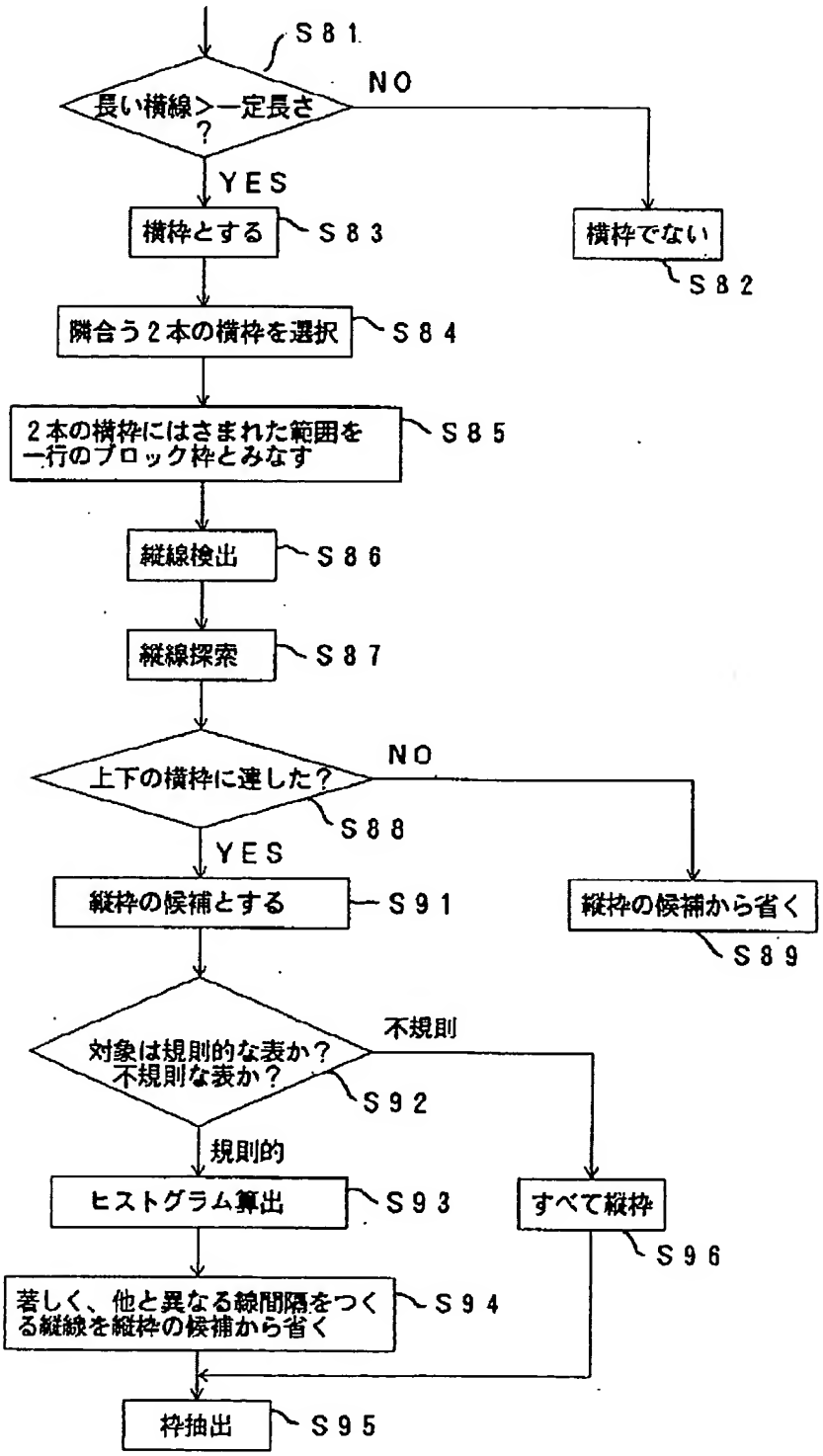




[Drawing 42]
部分パターンの水平方向と垂直方向の投影を示す図



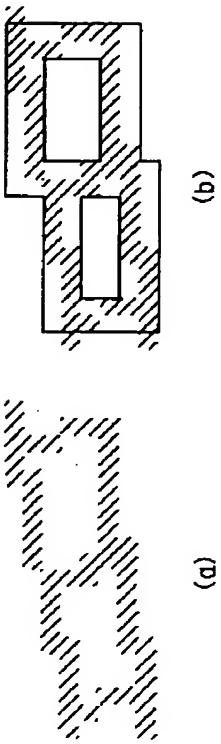

[Drawing 21]

枠検出部４３全体の処理を説明するフローチャート



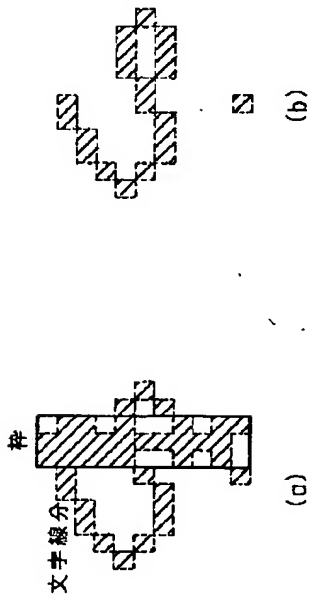

[Drawing 31]

枠が傾斜している場合の処理を説明する図



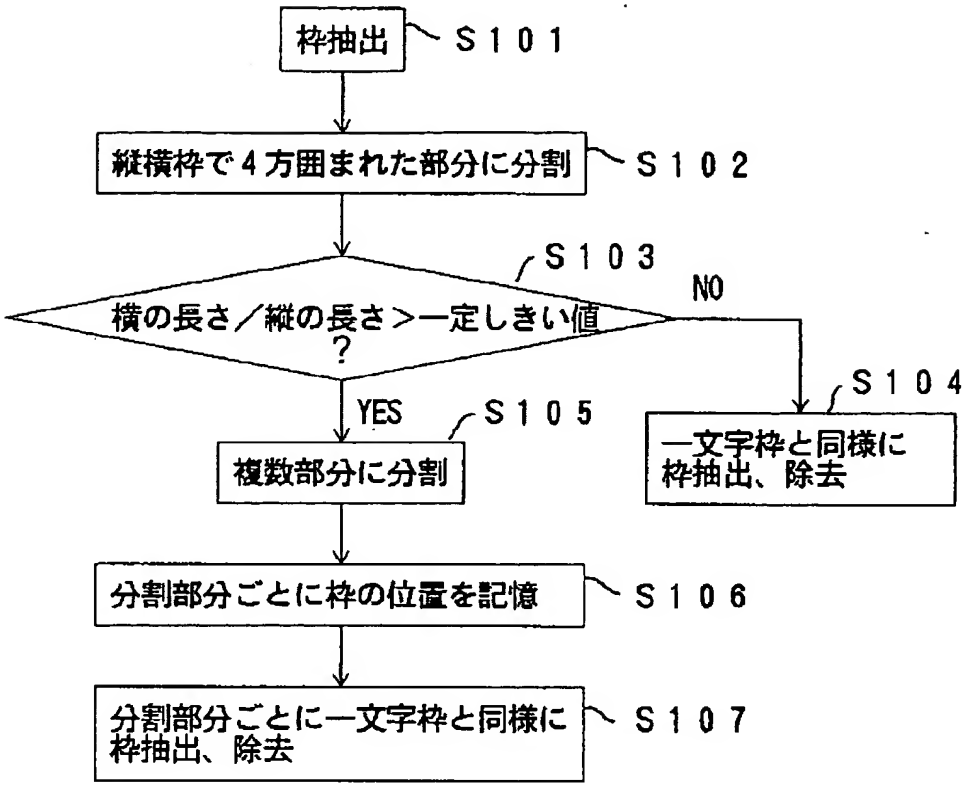

[Drawing 34]

9の誤った対応付けの例を説明する図



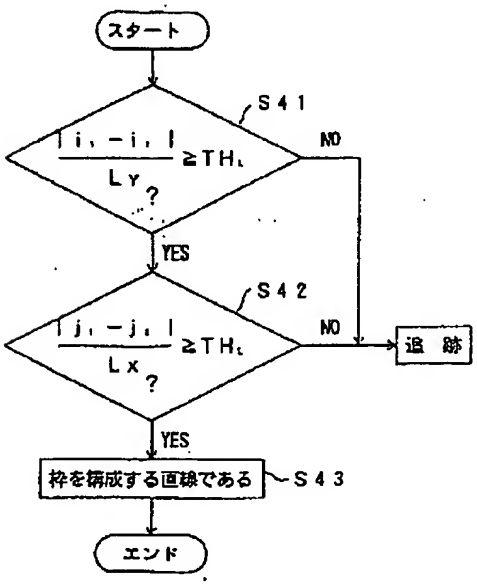

[Drawing 23]

枠分離部４４全体の処理を説明するフローチャート



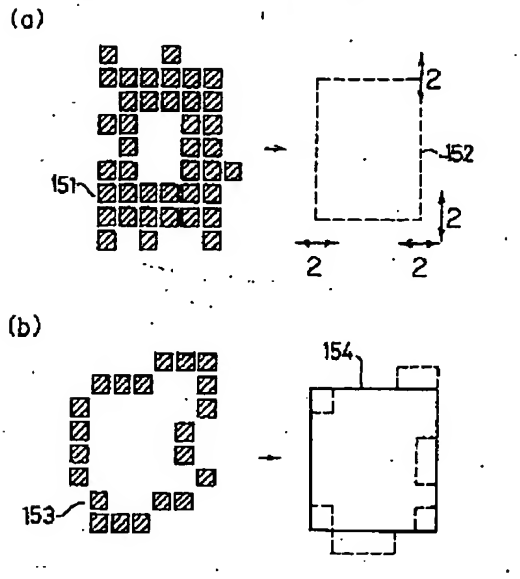

[Drawing 45]

４辺検出部２２ｃの処理を説明するフローチャート



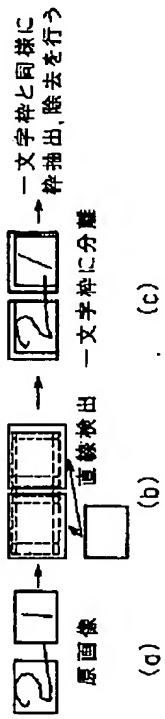

[Drawing 46]

線幅の太い枠および線幅の細い枠の一例を示す図




[Drawing 24]

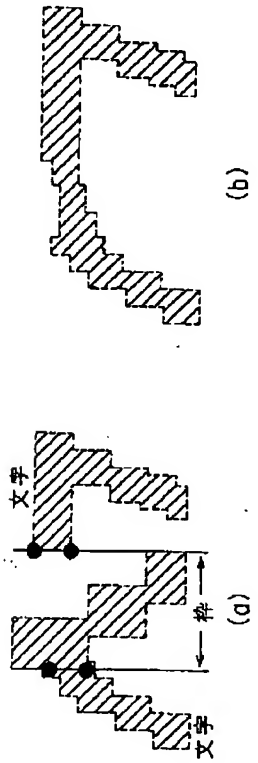
隣合う一文字枠が文字によって繋がっている場合の処理を説明する図




[Drawing 30]

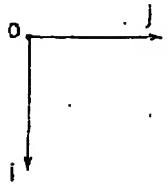
枠幅が文字幅よりかなり大きい場合の処理を
説明する図

[Drawing 33]
1と7とが誤って対応付けられる例を説明する図

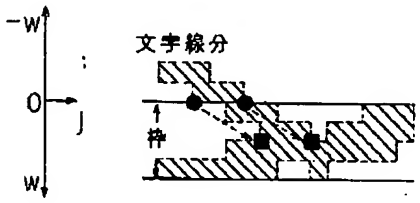



[Drawing 52]

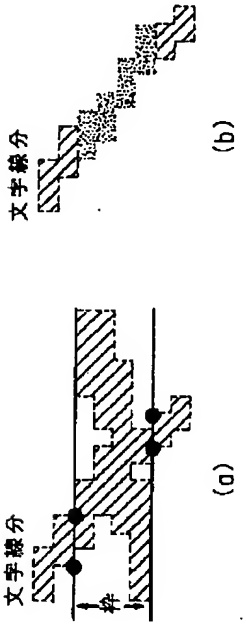
辺幅算出部240で用いる座標系を示す図




[Drawing 56]
枠内の交点の追跡例を示す図

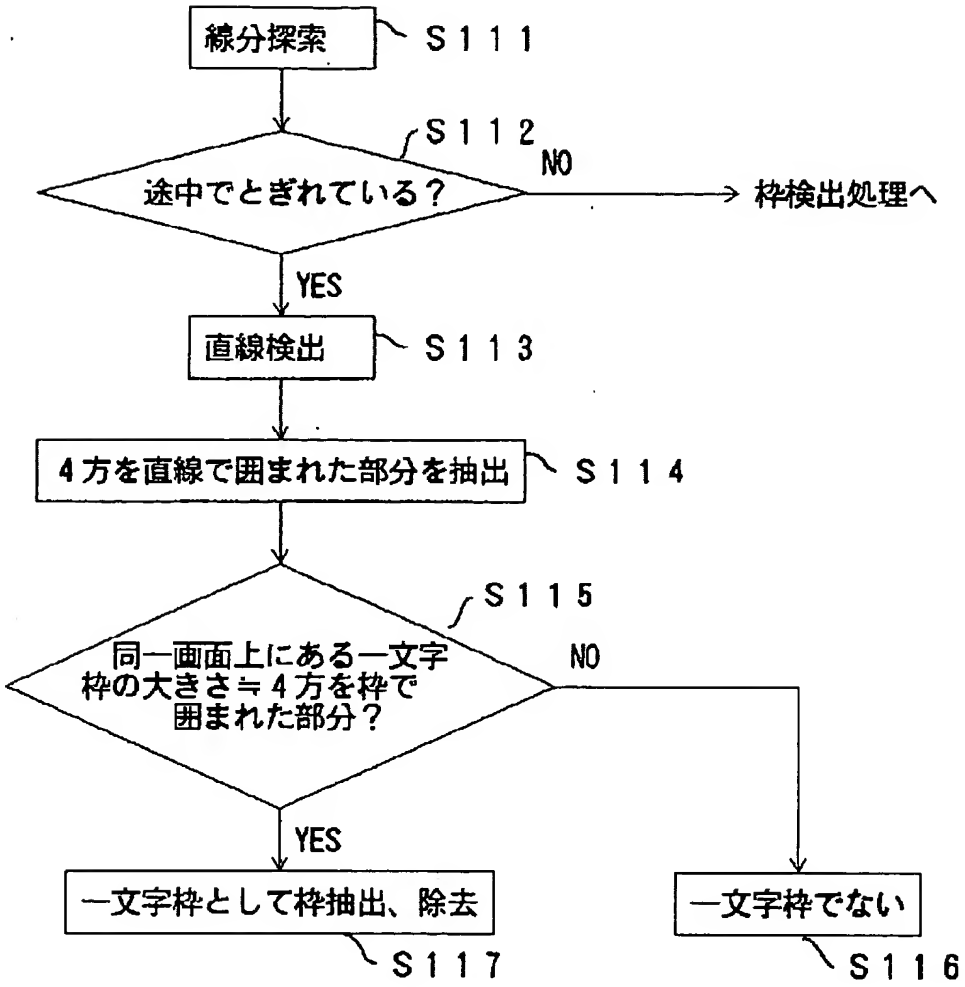



[Drawing 68]
枠内文字の抽出例を示す図

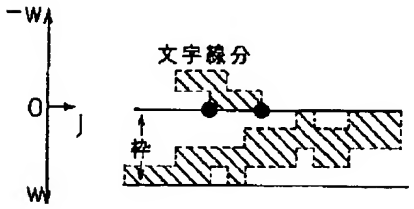



[Drawing 25]

枠抽出・除去部４５全体の処理を説明するフローチャート

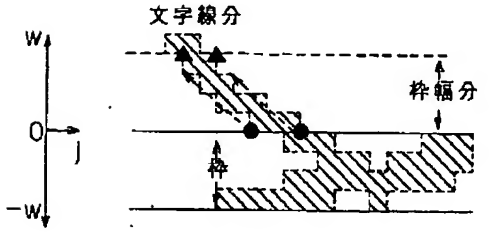



[Drawing 57]
枠内に交点が存在しない場合の処理を説明する図



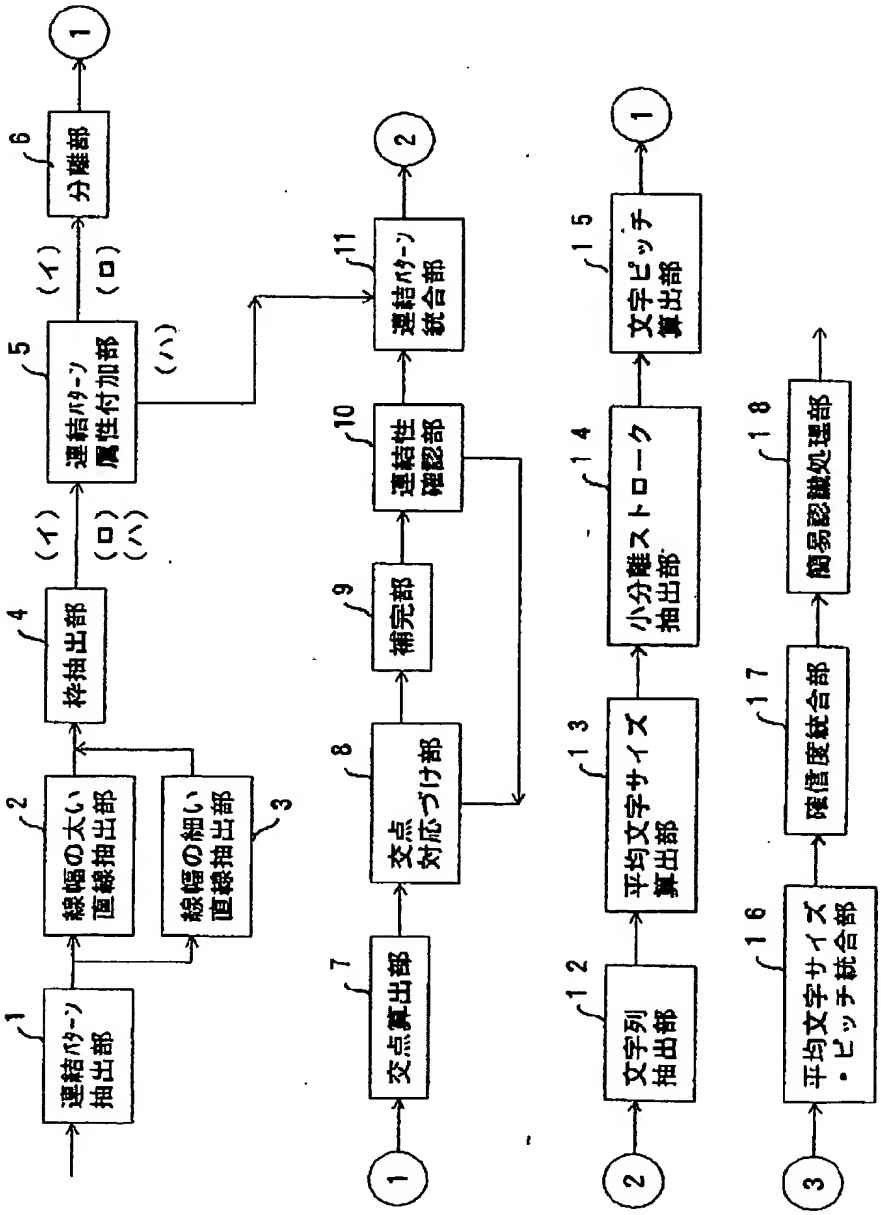

[Drawing 59]

枠から遠い交点の追跡例を示す図



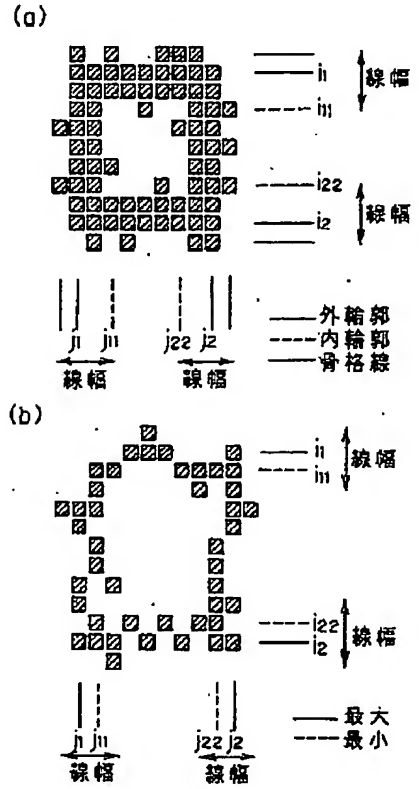

[Drawing 29]

先に提案された画像抽出方式を示すブロック図



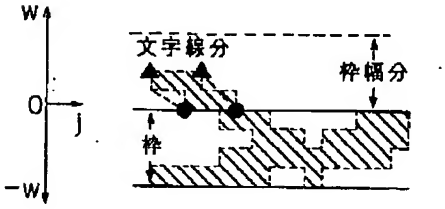
[Drawing 51]

線幅の太い枠と線幅の細い枠の骨格線、外輪郭等を示す図



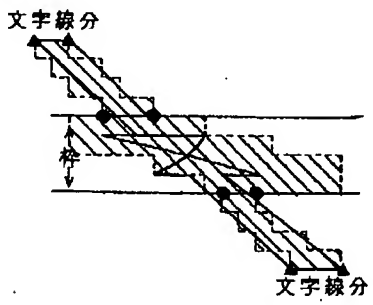

[Drawing 60]

追跡途中で文字線分がなくなる場合の処理を説明する図



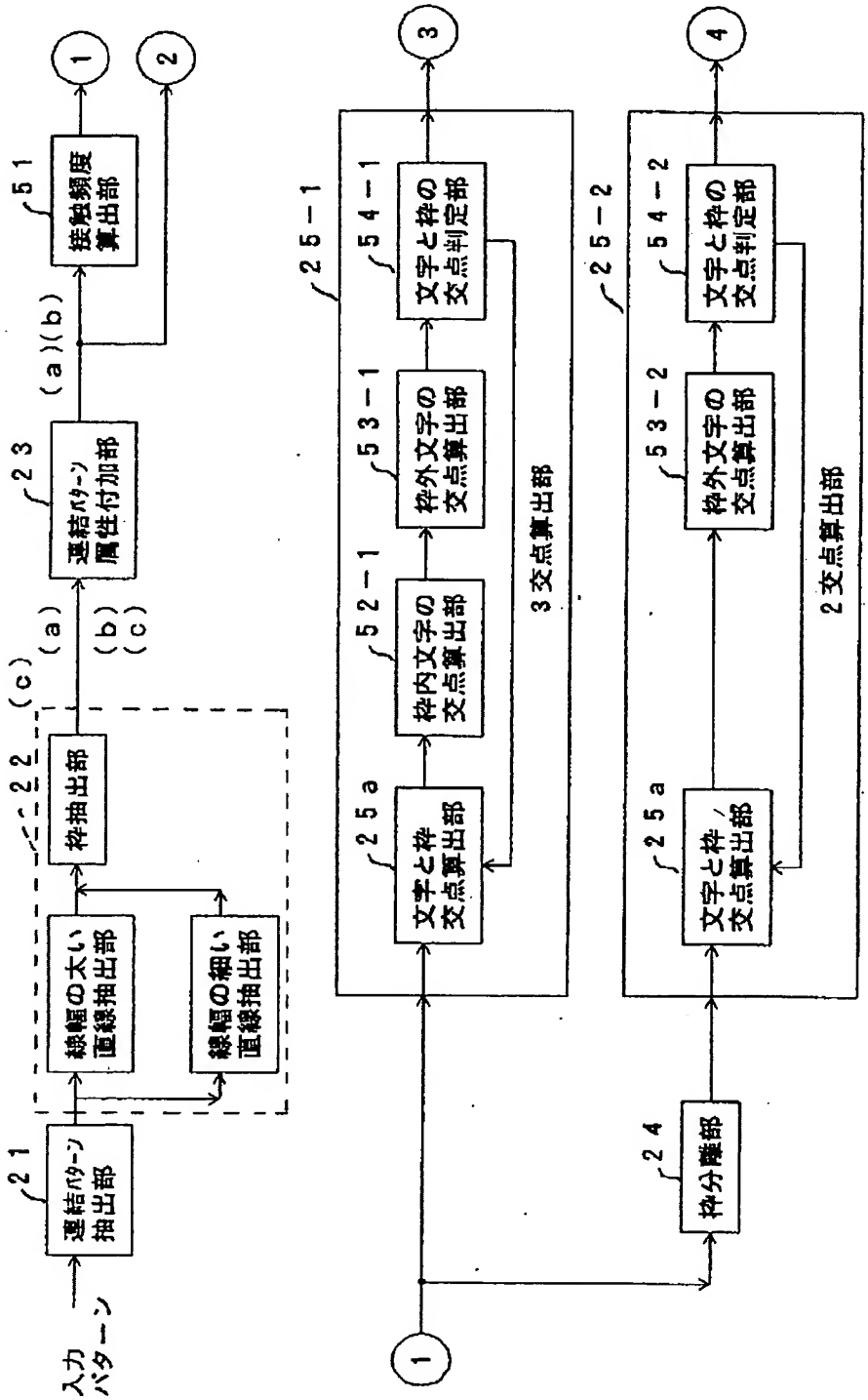

[Drawing 62]

方向性が一致する場合の交点の対応付けを示す図



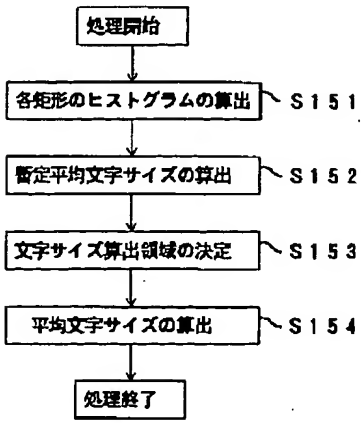

[Drawing 36]

本発明になる画像抽出装置の第２実施例を示すブロック図

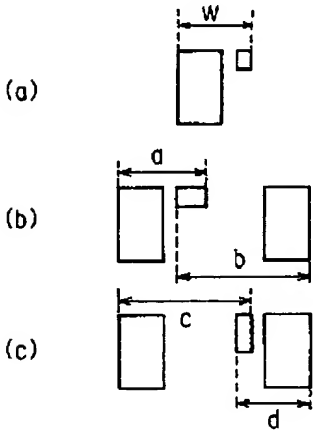



[Drawing 79]

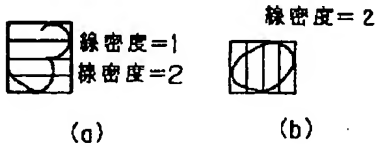
平均文字サイズ算出処理を説明するフローチャート



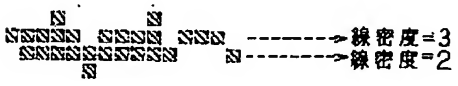

[Drawing 82]
統合を説明する図




[Drawing 87]
線密度の算出方法を説明する図



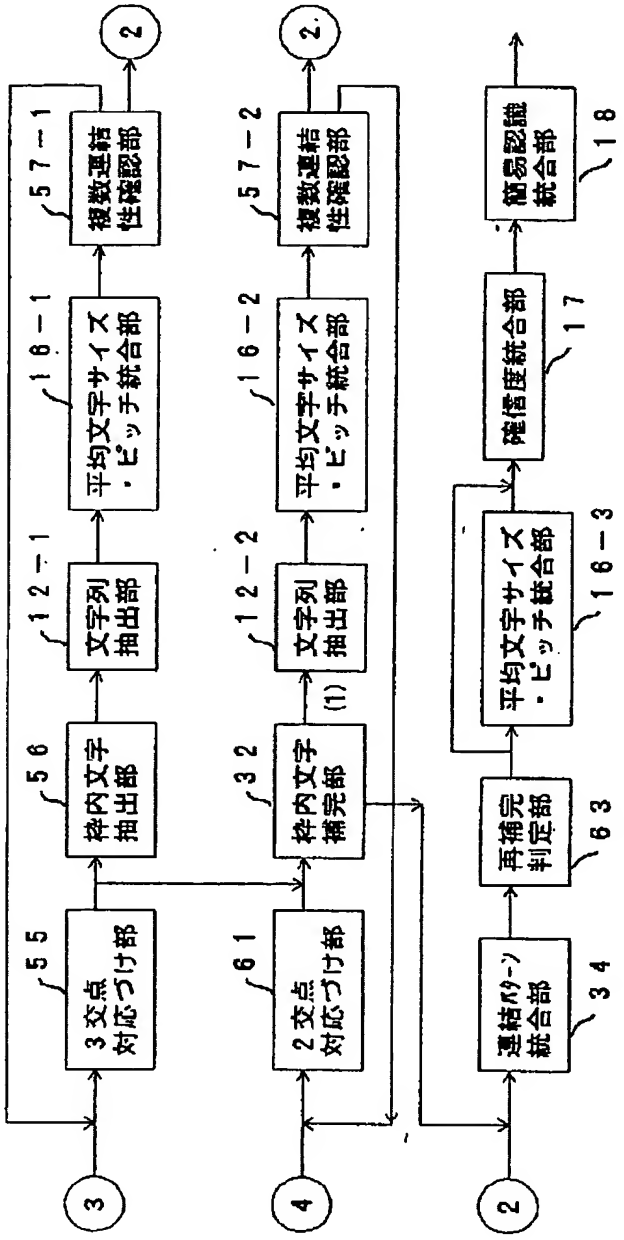

[Drawing 88]
横長のストロークに対して横方向線密度を
算出した場合の失敗例を示す図




[Drawing 91]
線密度の算出方法を説明する図

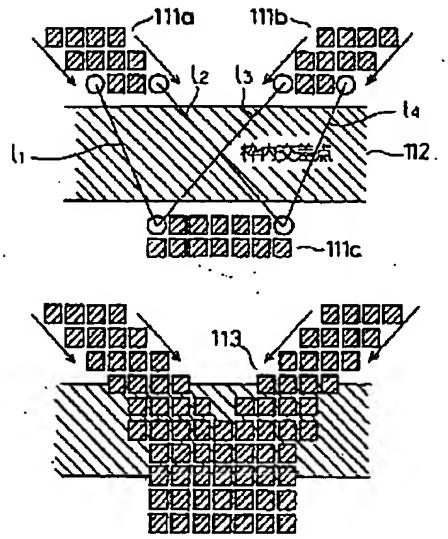
[Drawing 37]

本発明になる画像抽出装置の第２実施例を示すブロック図



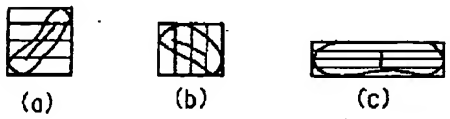

[Drawing 72]

交差点枠内補完の一例を示す図



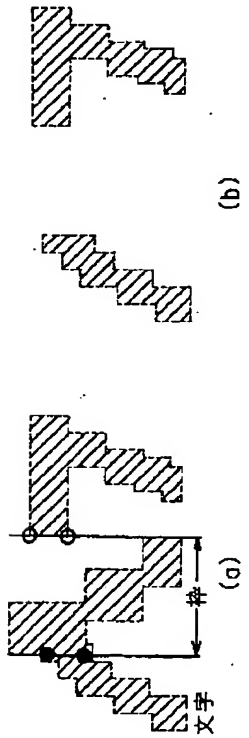

[Drawing 89]

傾きの算出方法を説明する図



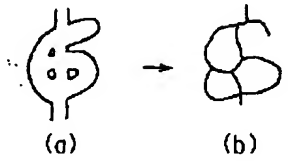

[Drawing 94]

第2実施例を図33の場合に適用した結果を説明する図



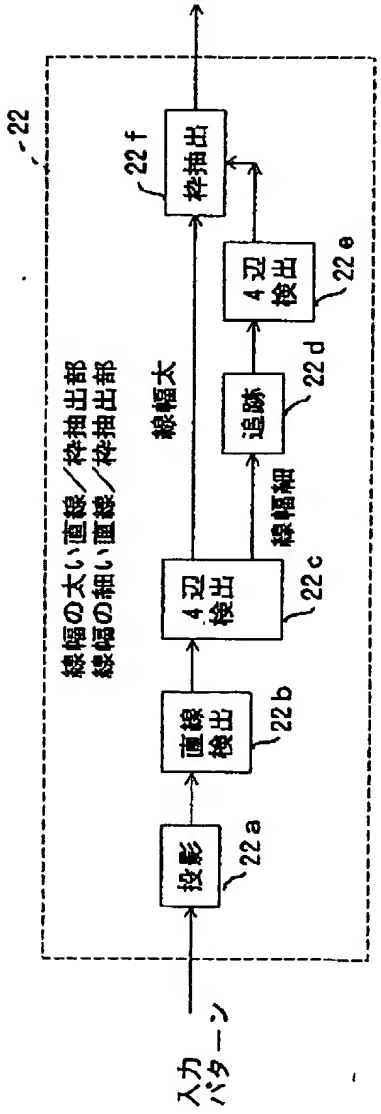

[Drawing 98]

細線化の問題を説明する図



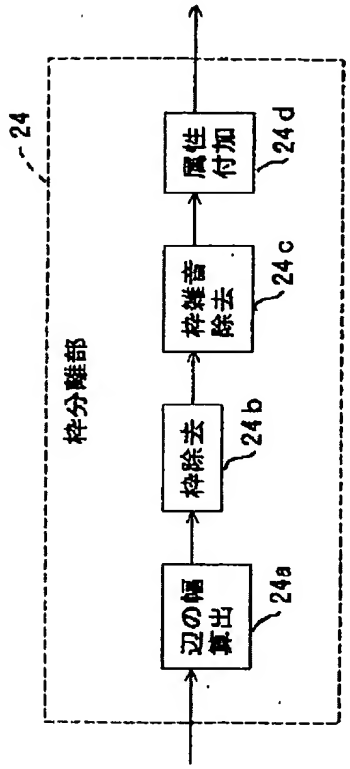

[Drawing 38]

直接／枠抽出部２'２の一実施例を示すブロック図



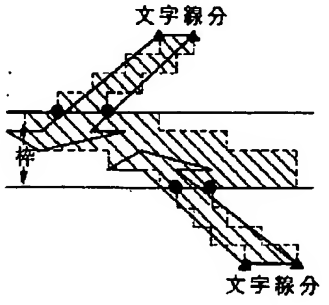

[Drawing 39]

枠分離部24の一実施例を示すブロック図



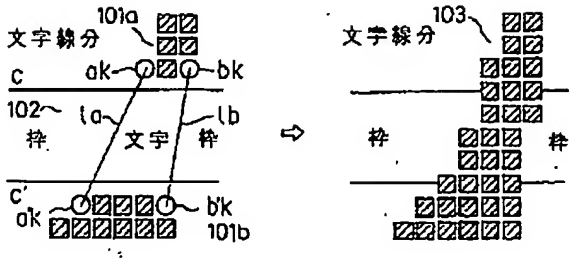

[Drawing 63]

方向性が一致しない場合の交点の対応付けを示す図



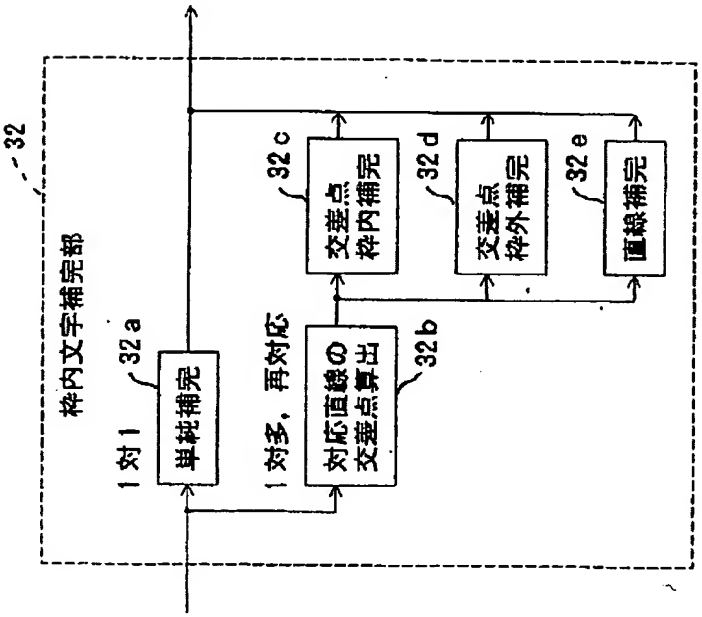

[Drawing 70]

一対一の単純補完の一例を示す図



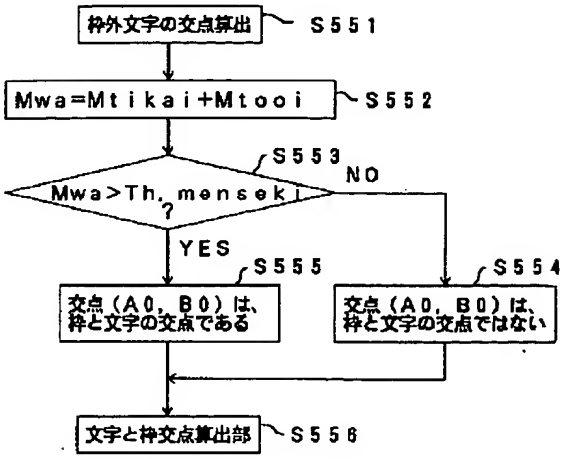

[Drawing 40]

枠内文字補完部３２の一実施例を示すブロック図

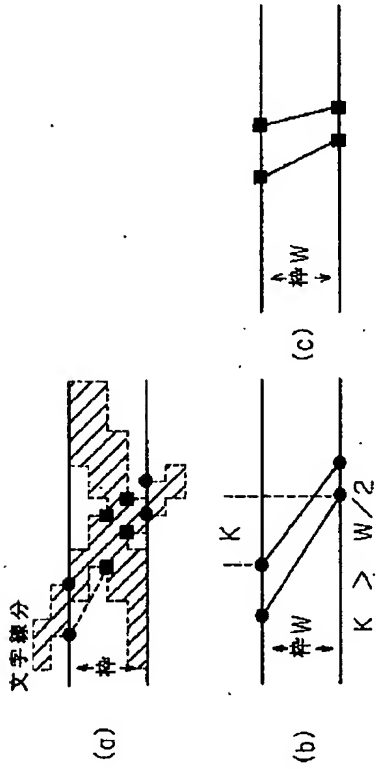



[Drawing 61]

文字／枠交点判定部５４－１又は５４－２の処理の
実施例を説明するフローチャート

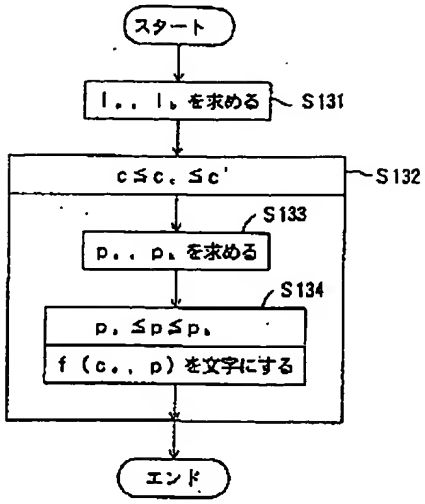



[Drawing 65]
枠内の交点により対応付けが行われる場合を
示す図



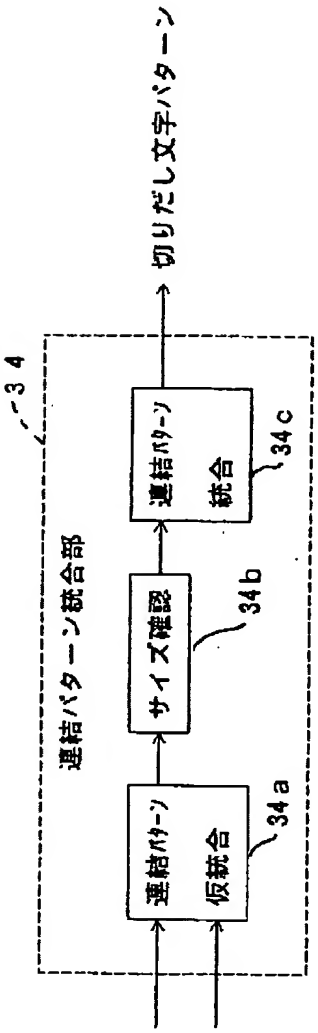

[Drawing 71]

単純補完部３２ａの処理を説明するフローチャート



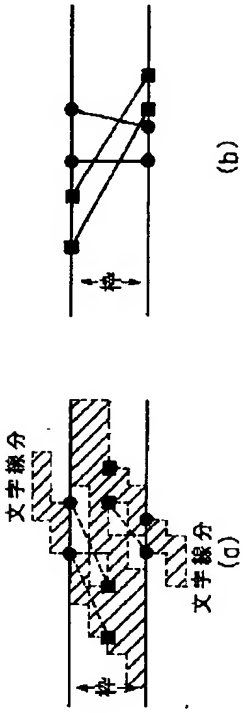

[Drawing 41]

連結パターン統合部３４の一実施例を示すブロック図



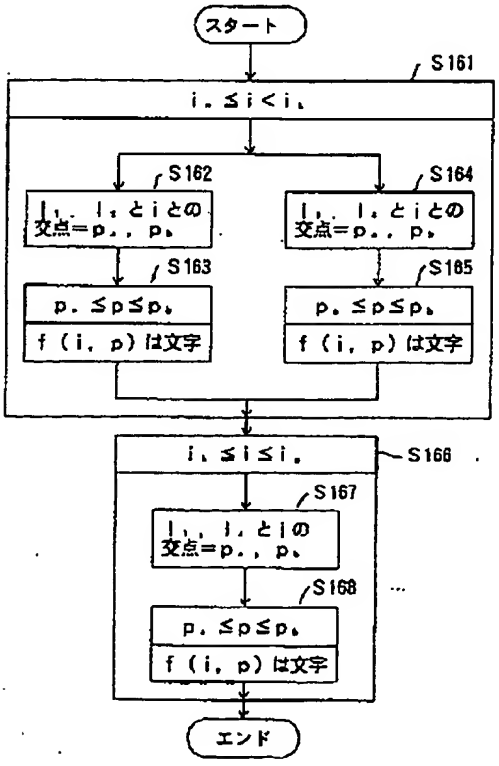

[Drawing 66]

枠内の交点により対応付けが行われない場合を示す図



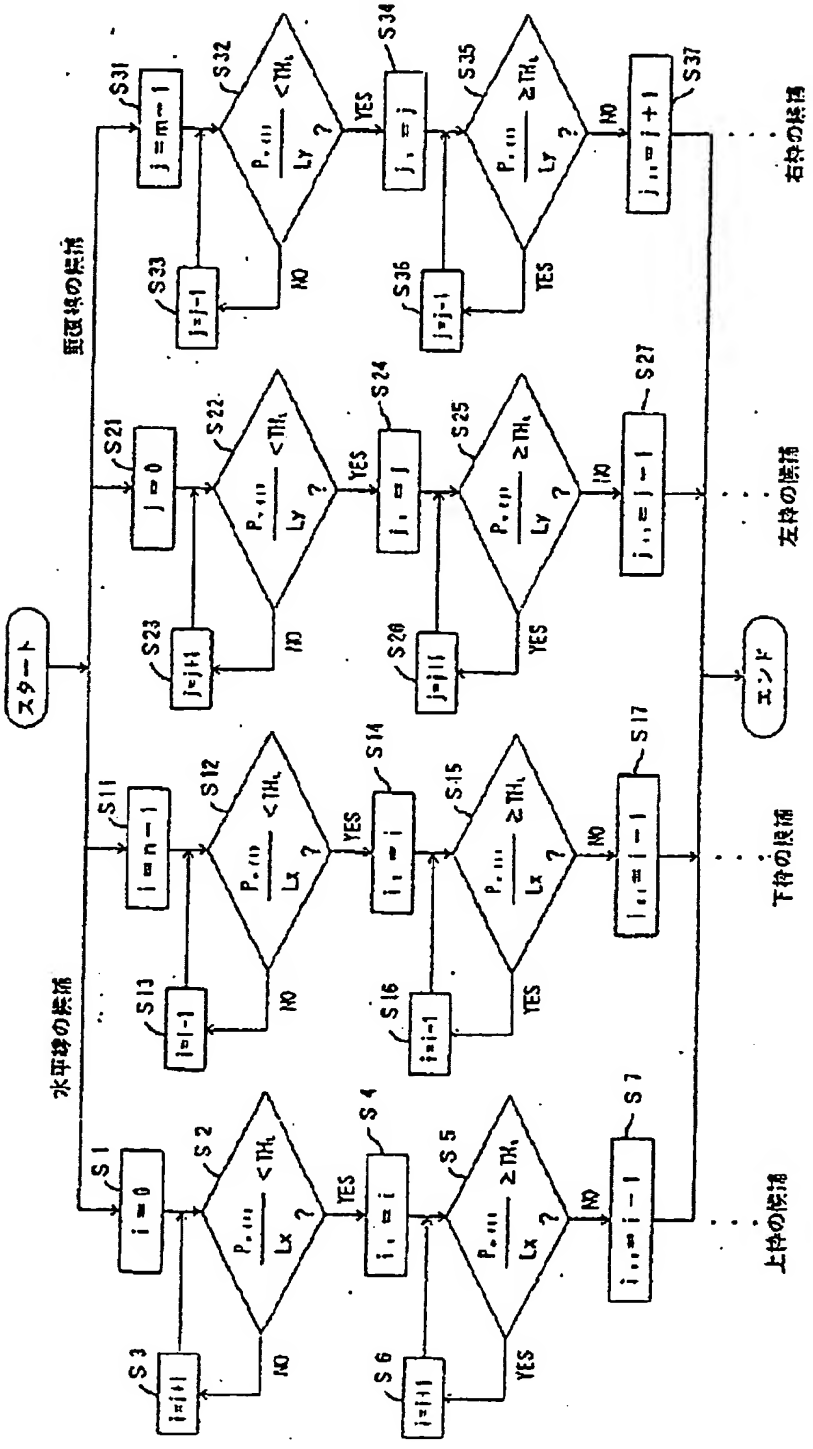

[Drawing 74]

交差点枠内補完部３２ｃの処理を説明するフローチャート




[Drawing 44]

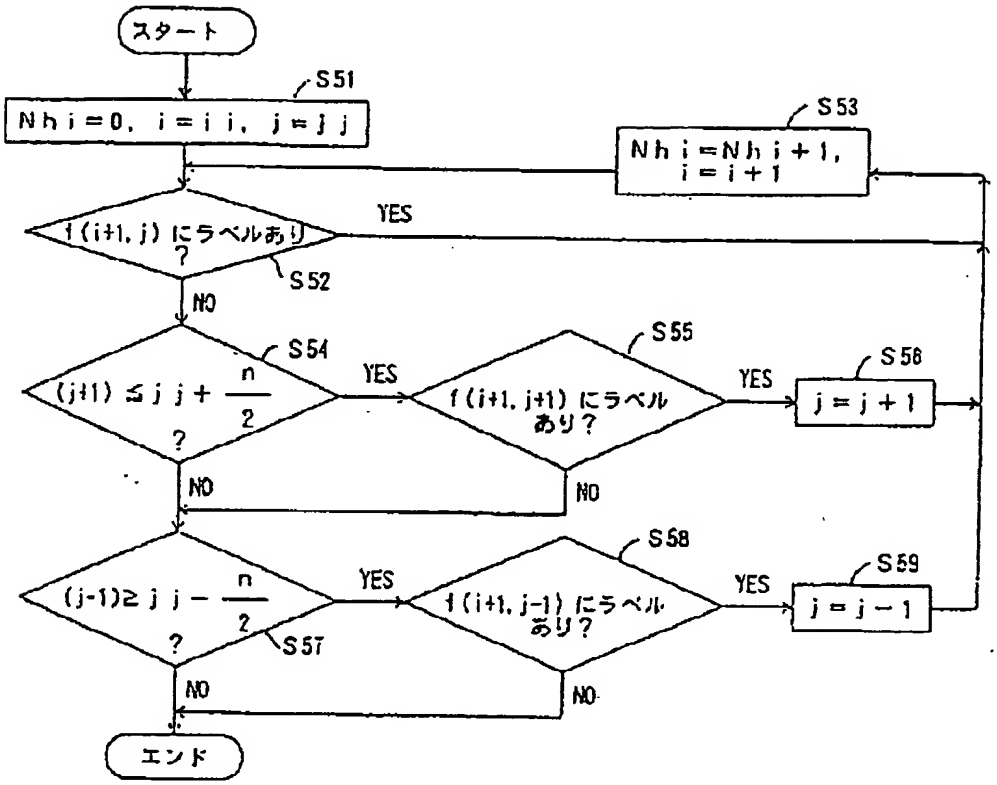
直線検出部２２ｂの処理を説明するフローチャート




[Drawing 48]

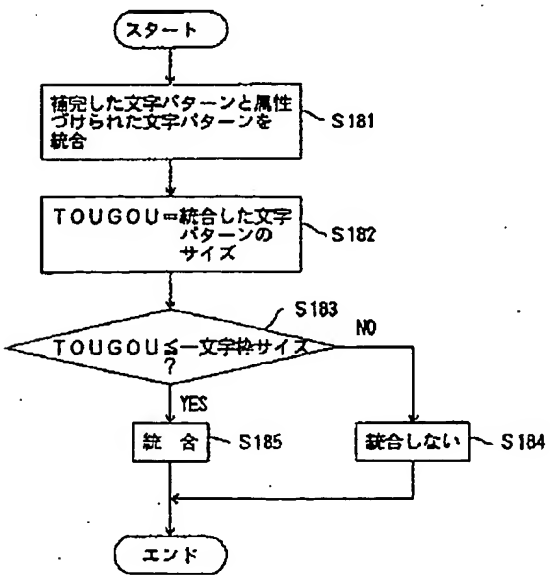
水平方向のｎラインランレングスを求める処理

を説明するフローチャート



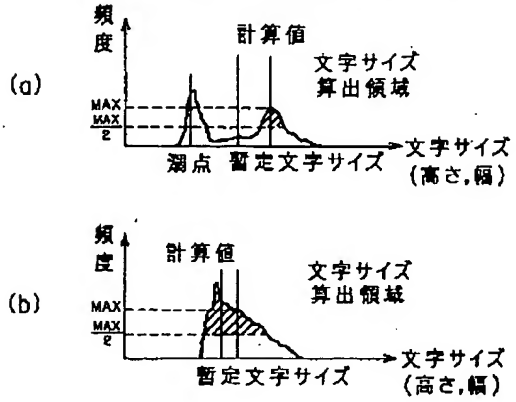

[Drawing 78]

連結パターン統合部３４の処理を説明するフローチャート



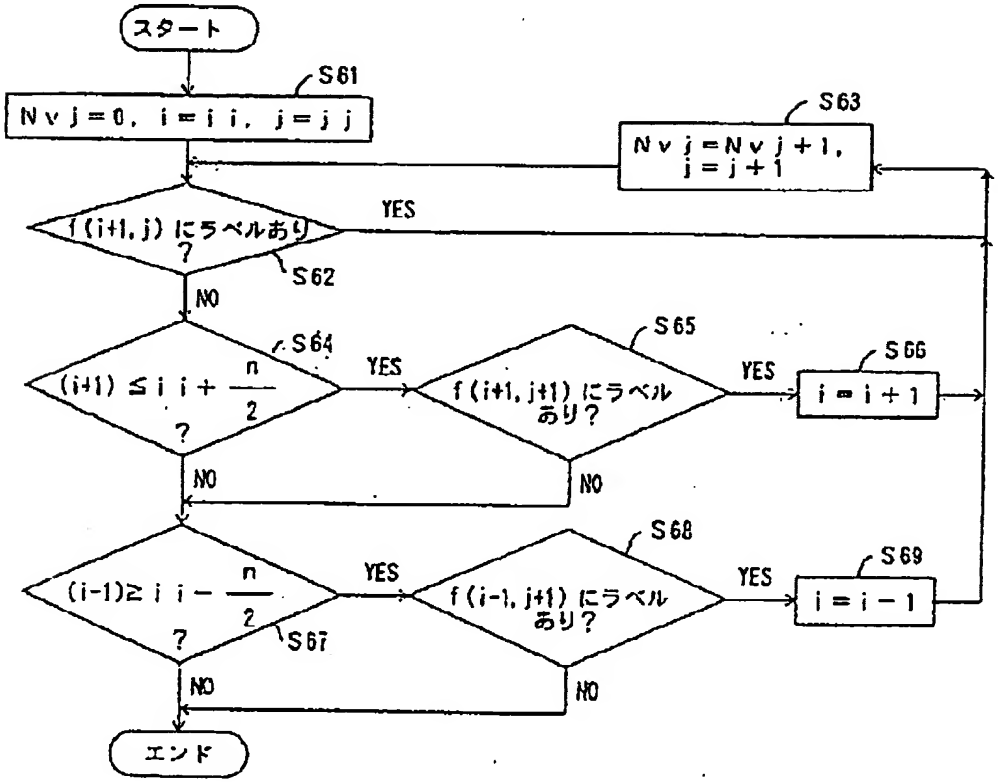

[Drawing 80]

平均文字サイズ算出方法を説明する図



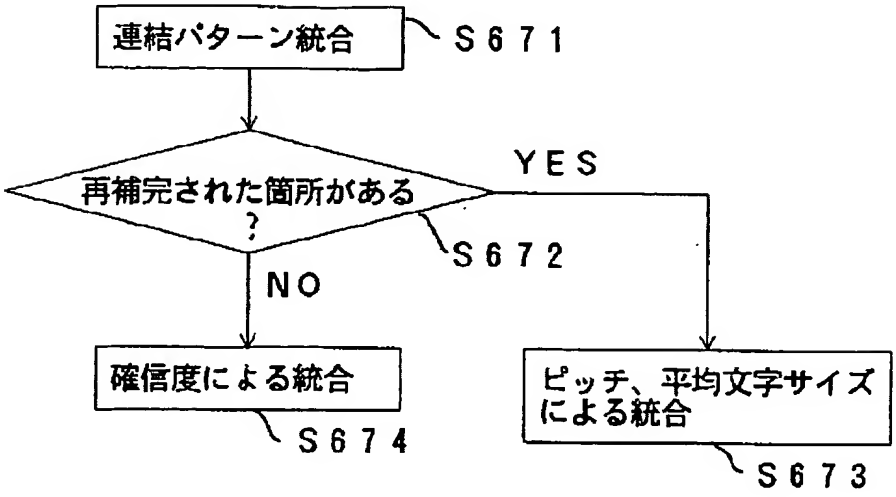

[Drawing 49]

垂直方向のnラインランレングスを求める処理を説明するフローチャート



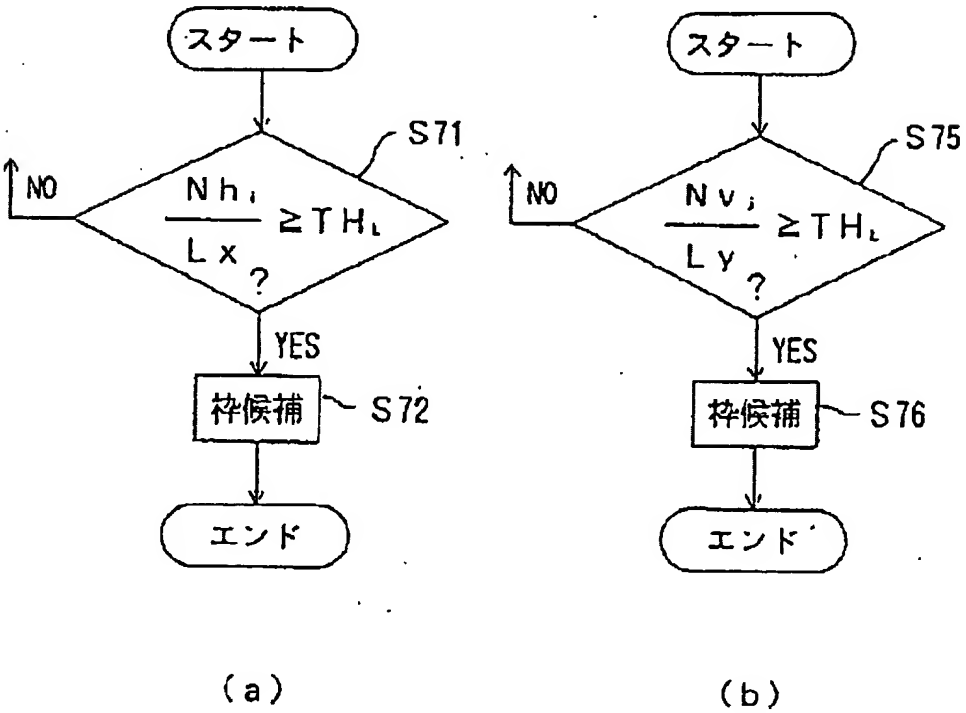

[Drawing 85]

再補完判定部６３の処理の一実施例を説明するフローチャート




[Drawing 50]

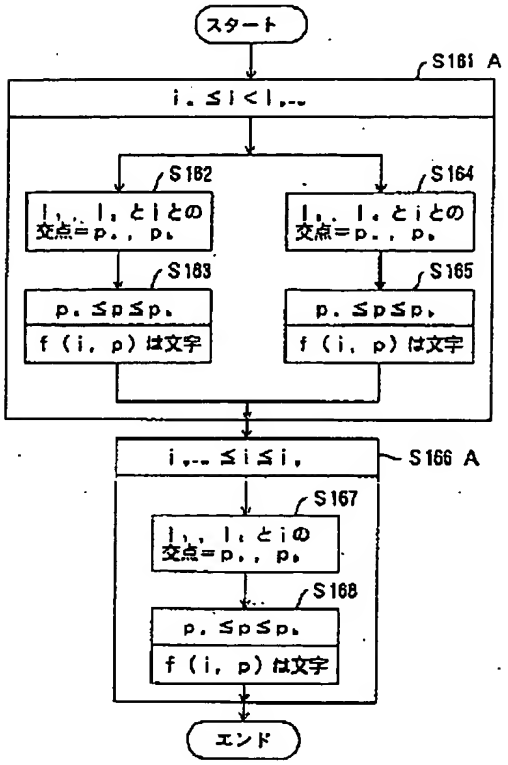
４辺検出部２２ｅの処理を説明するフローチャート




（ａ）　（ｂ）

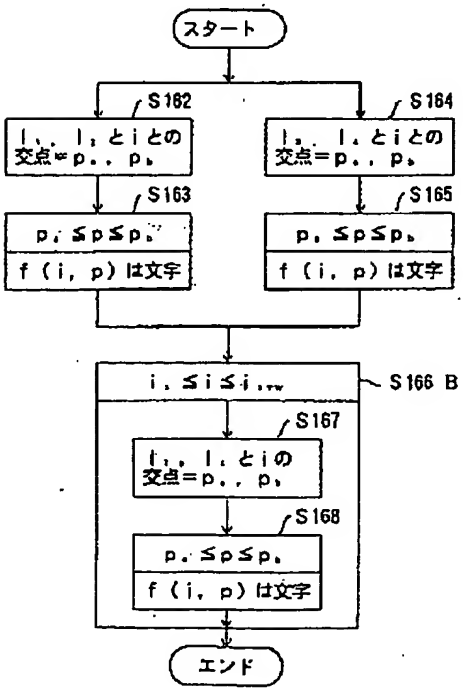
[Drawing 75]

交差点枠外補完部３２ｄの処理を説明するフローチャート



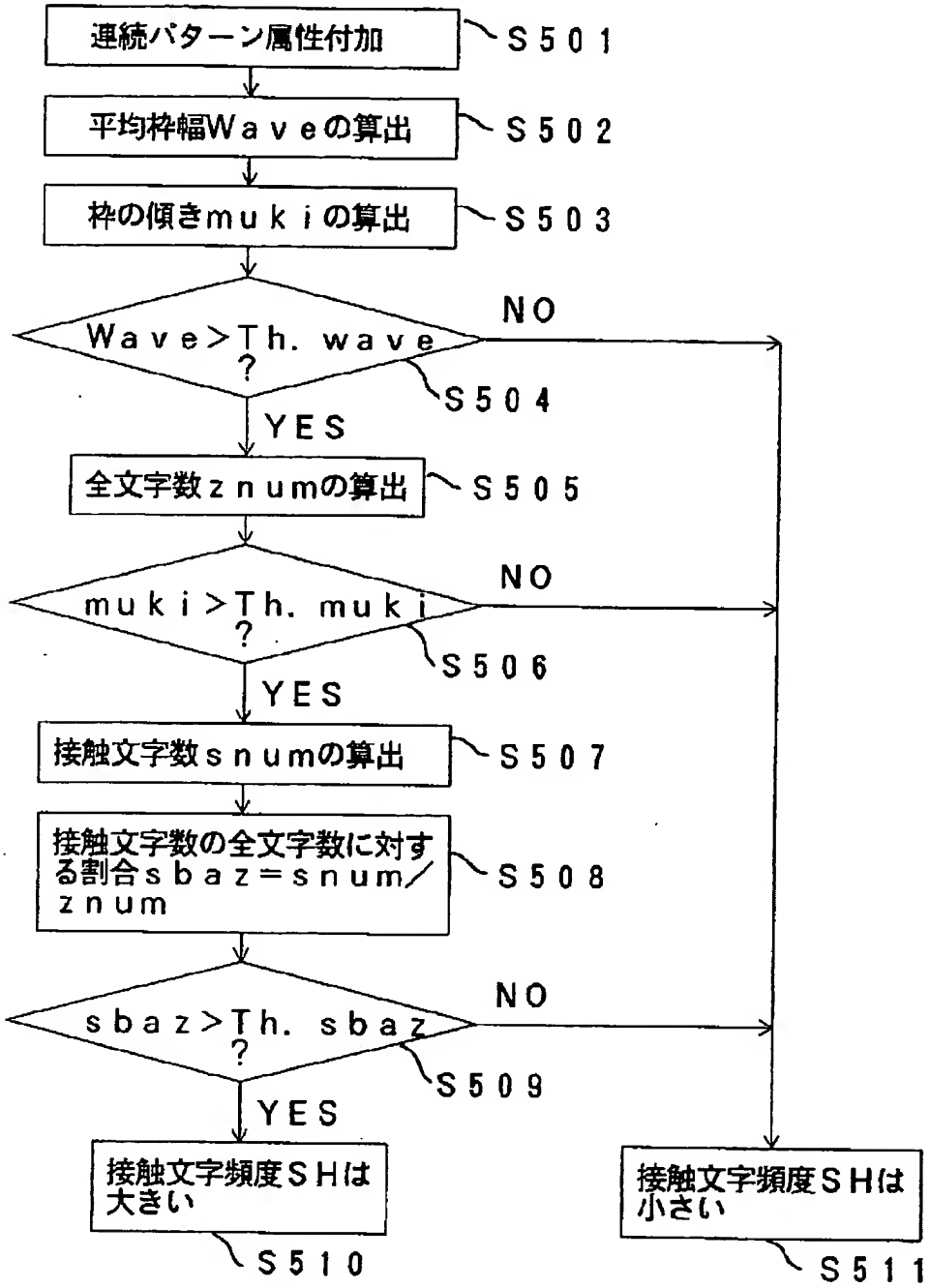

[Drawing 76]

直線補完部３２ｅの処理を説明するフローチャート



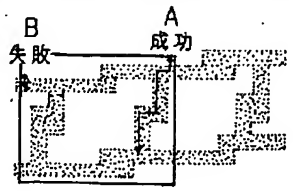

[Drawing 53]

接触頻度算出部５１の処理の一実施例を示すフローチャート

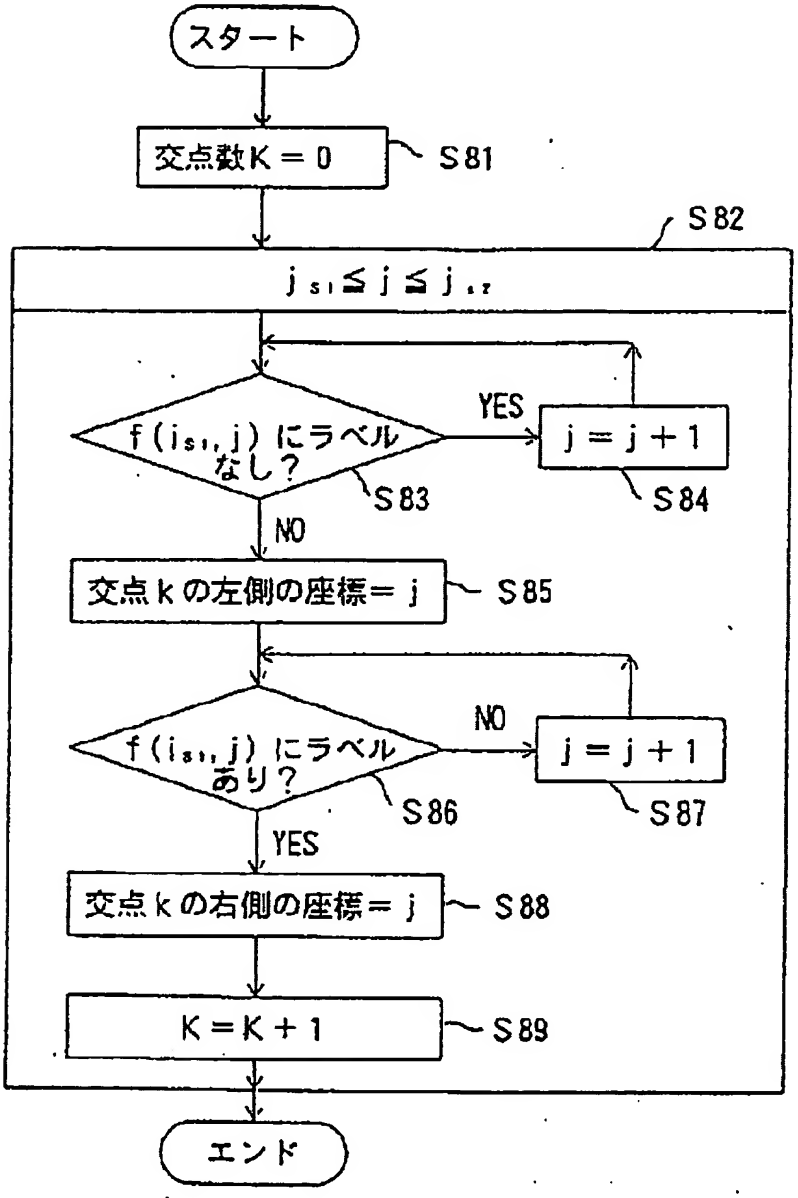



[Drawing 99]

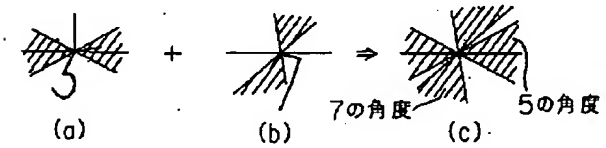
線分の探索の問題を説明する図



[Drawing 54]
交点算出部２５ａの処理を説明するフローチャート



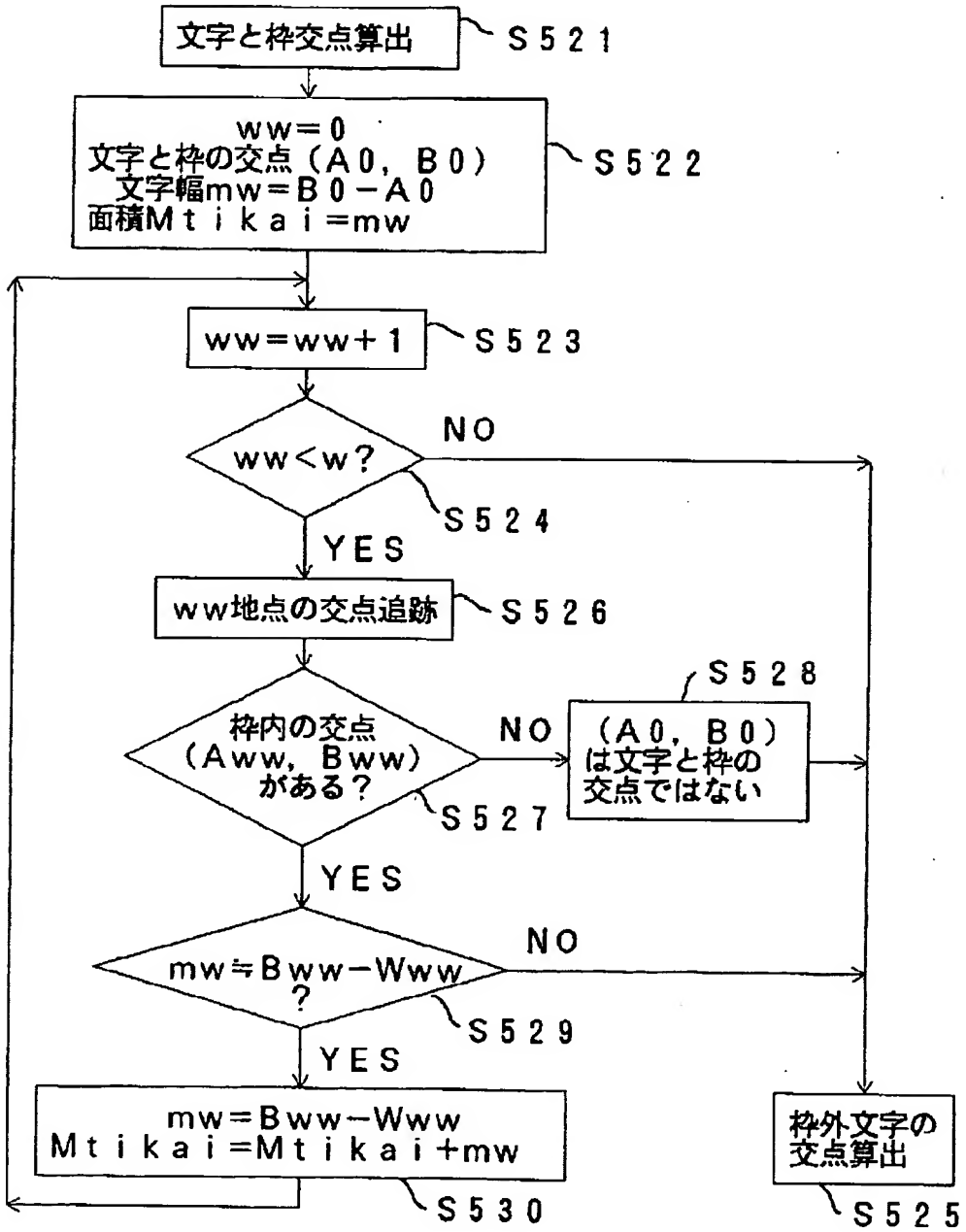

[Drawing 90]
5及び7の分離ストロークの角度を示す図




[Drawing 55]

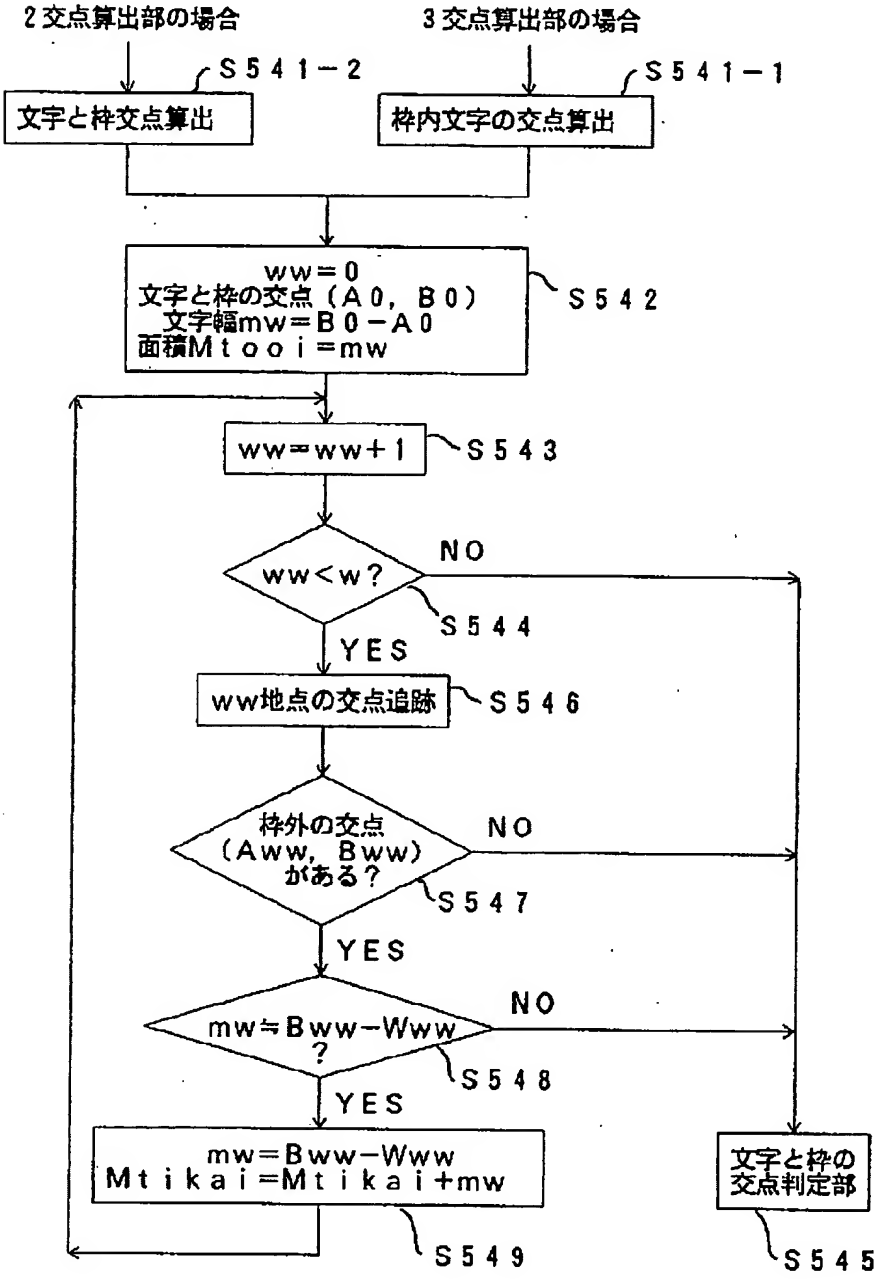
枠内文字の交点算出部５２－１の処理の

一実施例を説明するフローチャート



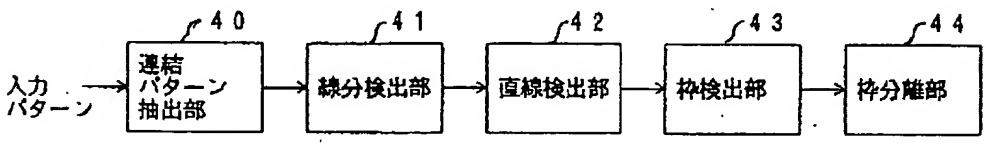

[Drawing 58]

枠外文字の交点算出部２５－１の処理の一実施例を説明するフローチャート



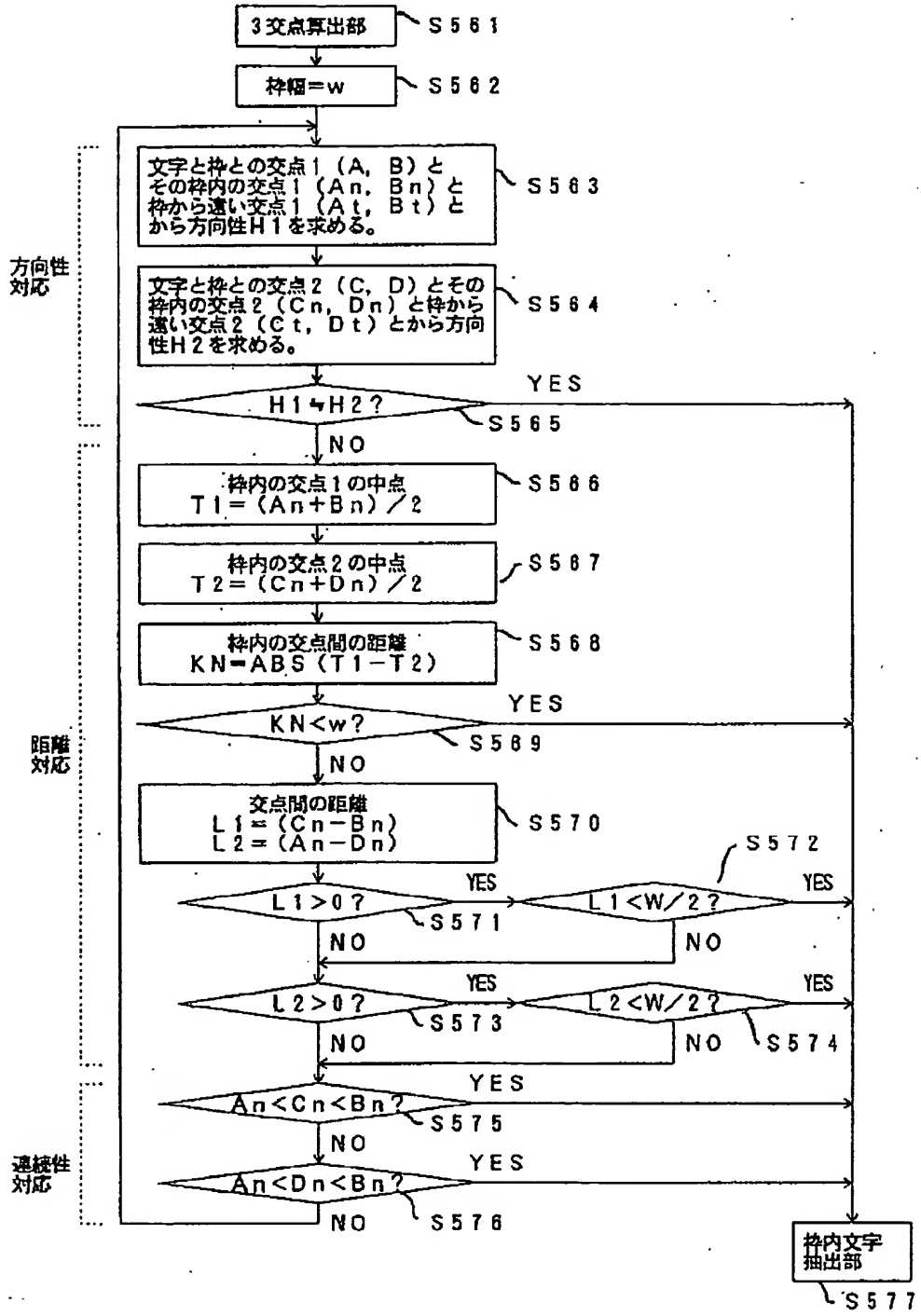

[Drawing 97]

先に提案されている画像抽出方式を示すブロック図



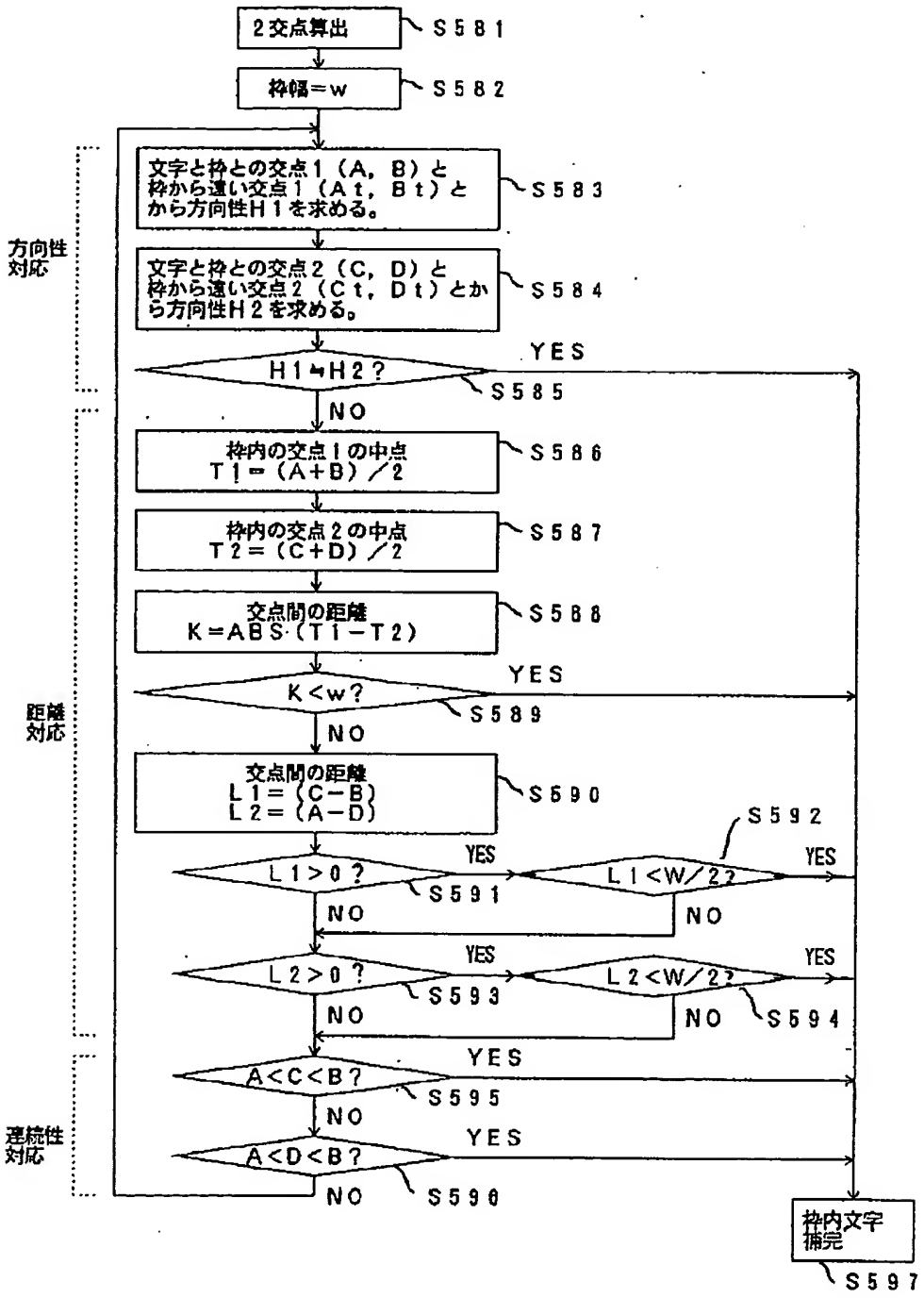

[Drawing 64]

３交点対応付け部５５の処理の一実施例を説明するフローチャート



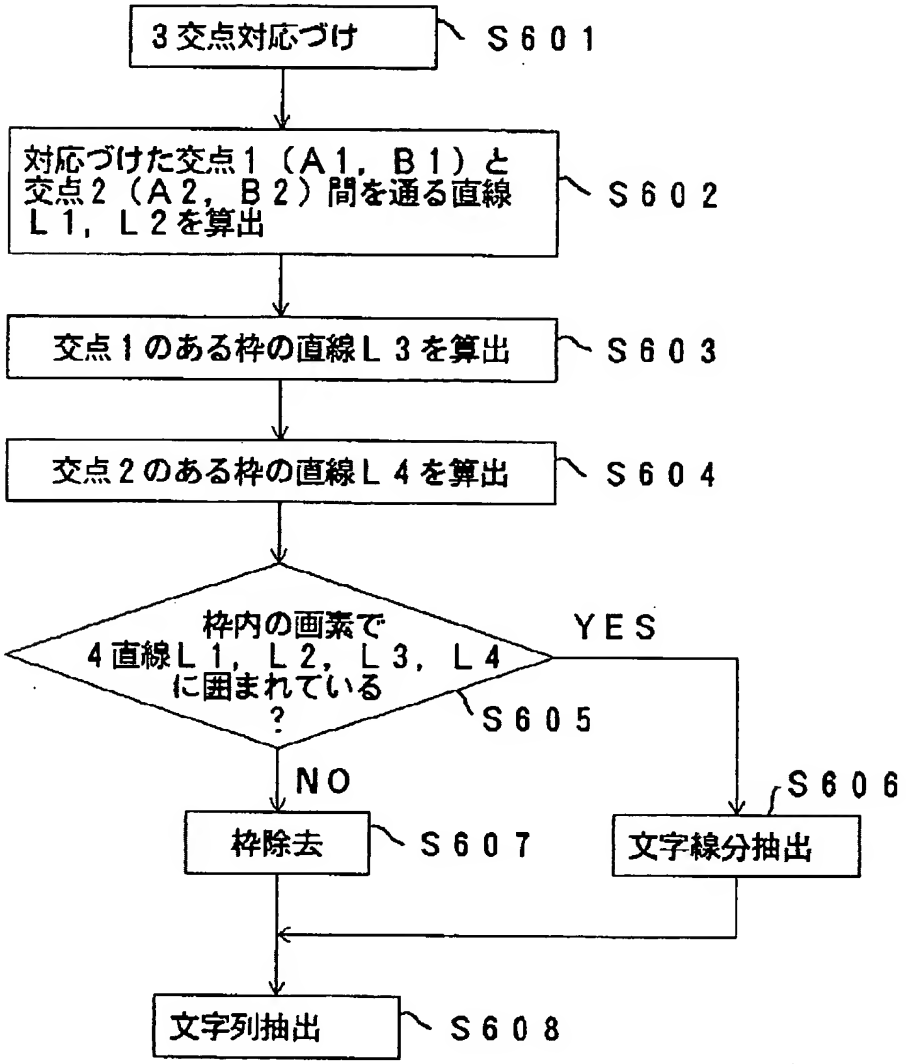

[Drawing 67]

２交点対応付け部６１の処理の一実施例を説明するフローチャート



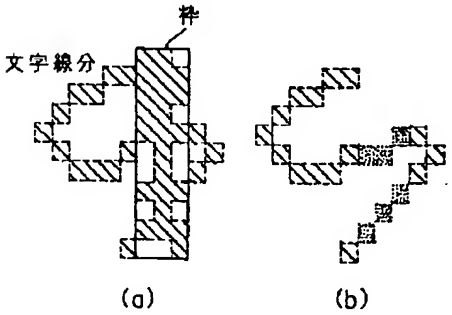

[Drawing 69]

枠内文字抽出部５６の処理の一実施例を説明するフローチャート



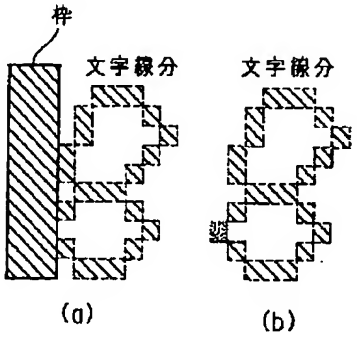

[Drawing 95]

第２実施例を図３４の場合に適用した結果を説明する図



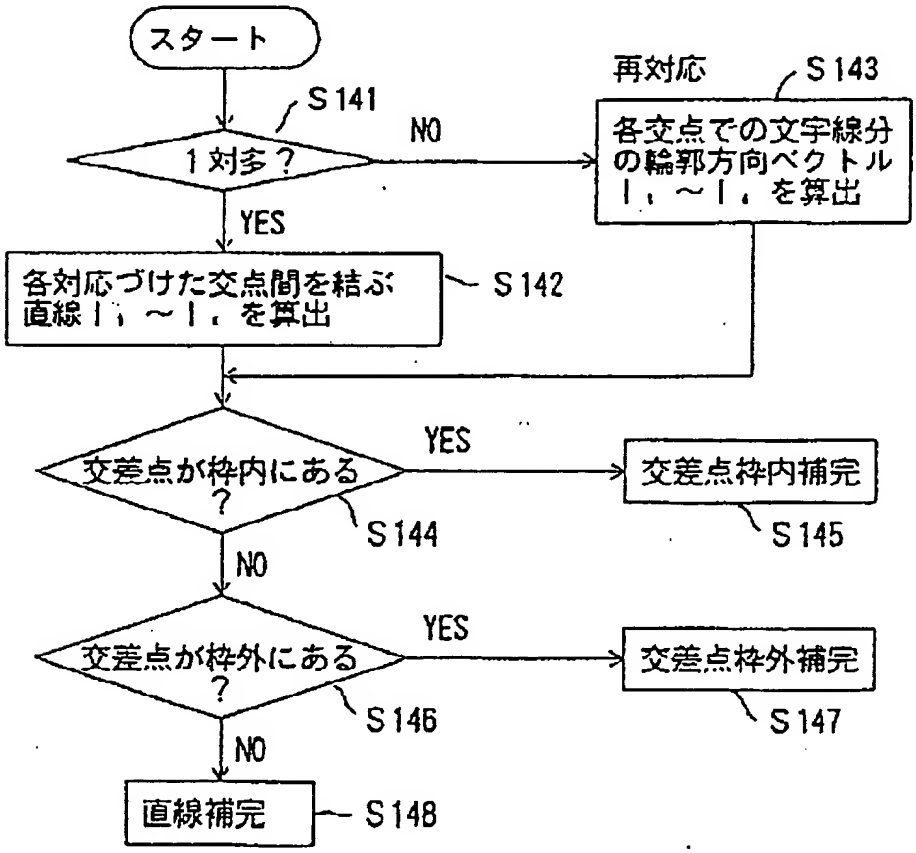

[Drawing 96]

第2実施例を図35の場合に適用した結果を説明する図



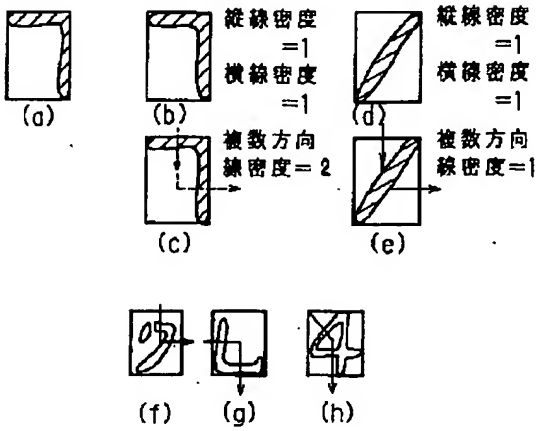

[Drawing 73]

対応直線の交差点算出部３２ｂの処理を説明するフローチャート



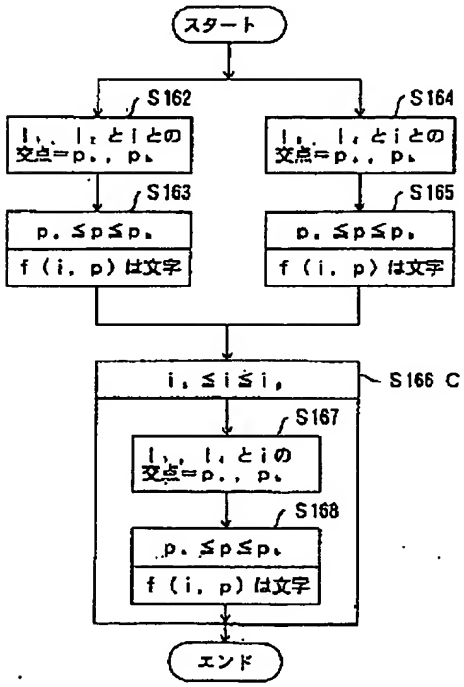

[Drawing 92]

複数方向の線密度の算出方法を説明する図



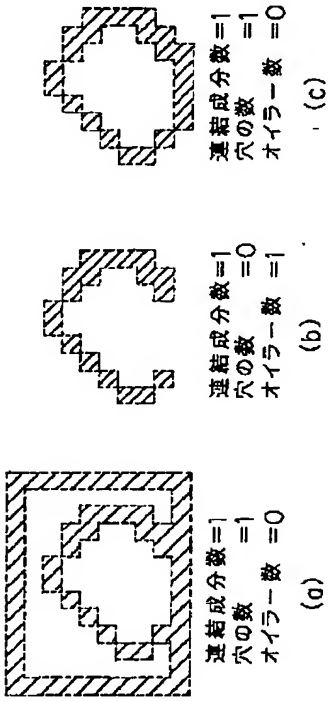

[Drawing 77]

対応付けが不可能な交点に対する処理を
説明するフローチャート



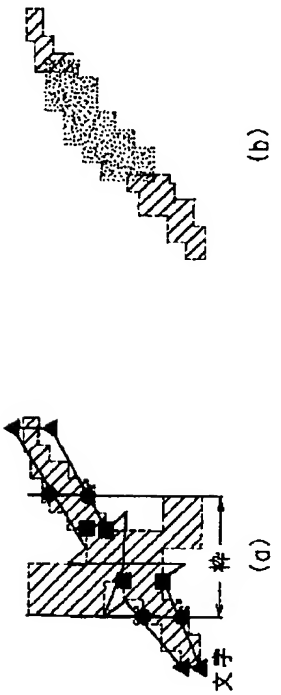

[Drawing 83]

原画像と処理された画像とで連結成分の数は変化しないが穴の数及びオイラー数が変化した場合を示す図



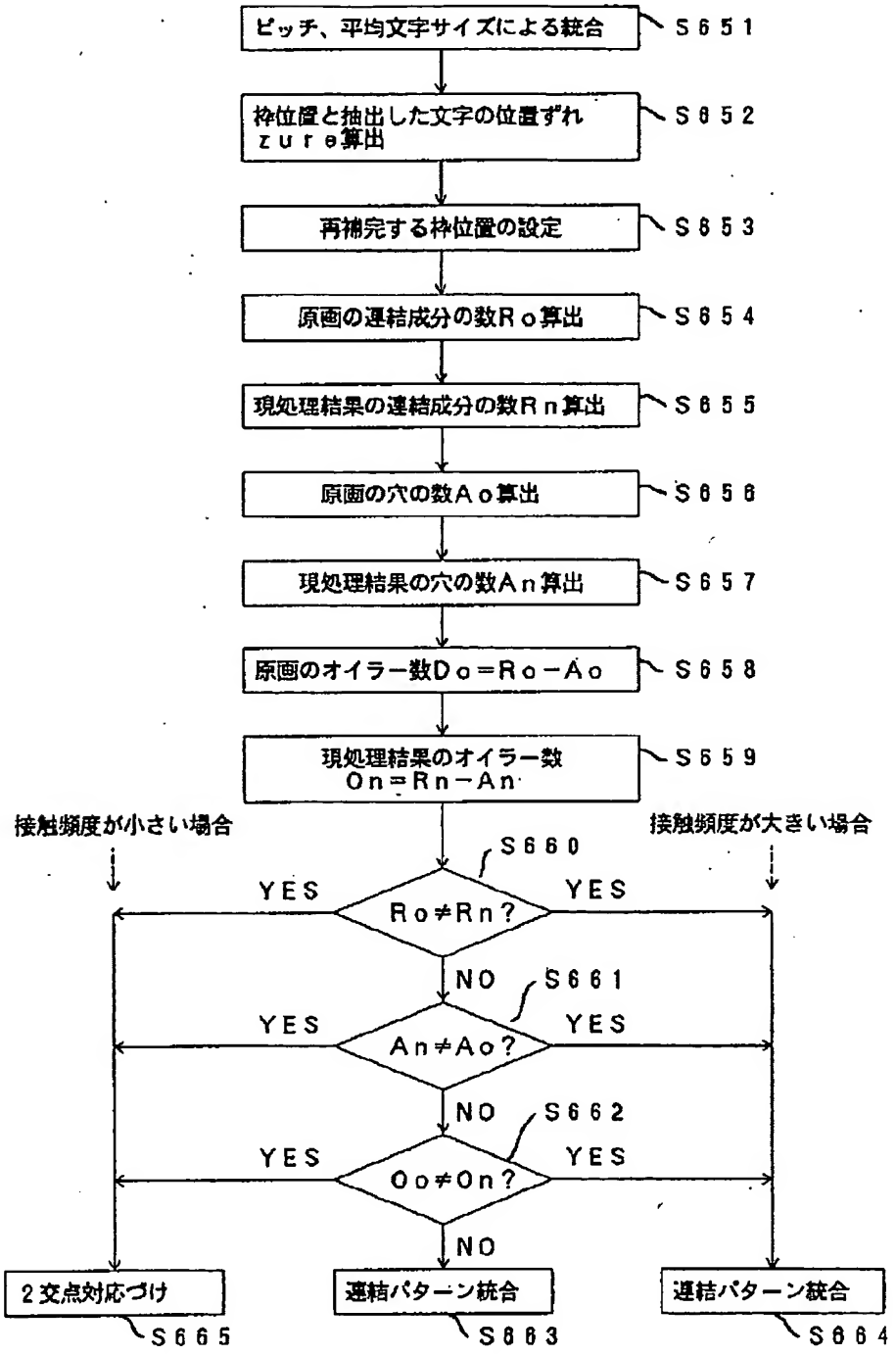

[Drawing 93]

第2実施例を図32の場合に適用した結果を説明する図



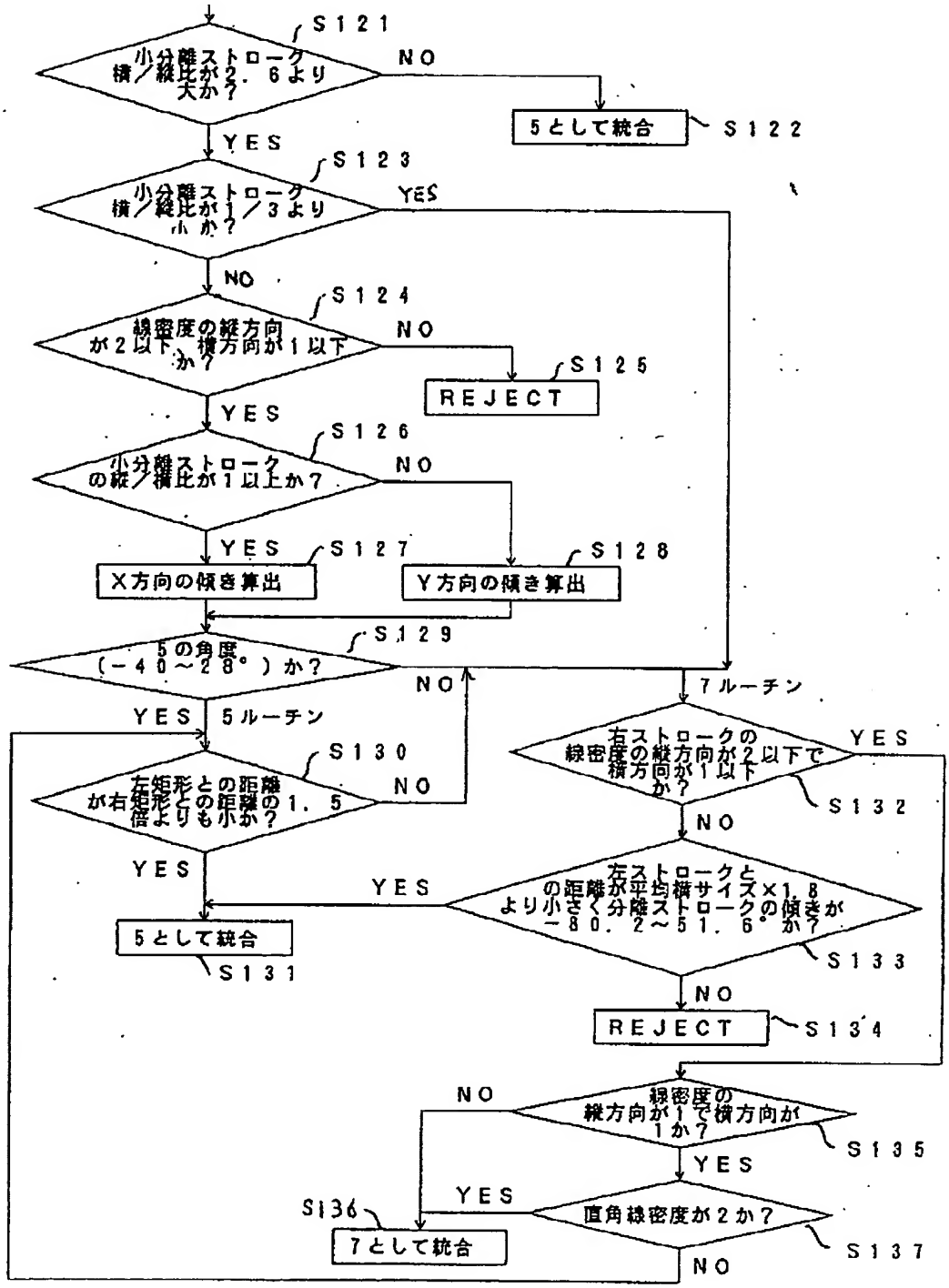

[Drawing 84]

複数連続性確認部５７－１の処理の一実施例を説明するフローチャート




[Drawing 86]

関易認識処理部１８の処理の一実施例を説明するフローチャート




[Translation done.]

* NOTICES *

**JPO and NCIPI are not responsible for any damages caused by the use of this translation.**

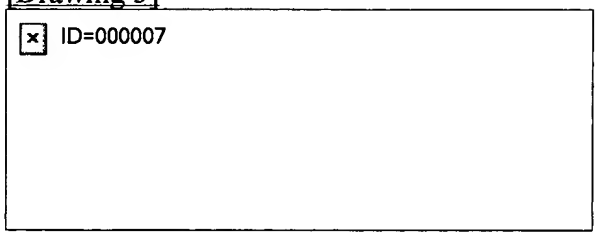
1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

---

DRAWINGS

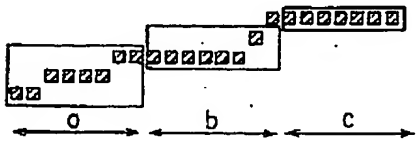
---

[Drawing 3]



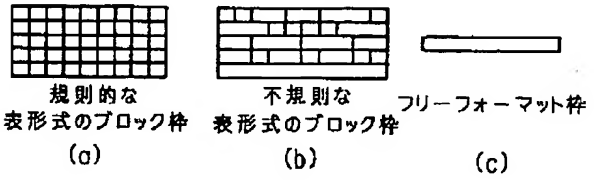

[Drawing 5]

矩形線分の検出を説明する図



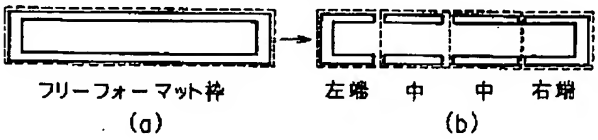

[Drawing 18]

処理の対象となる枠の種類を示す図



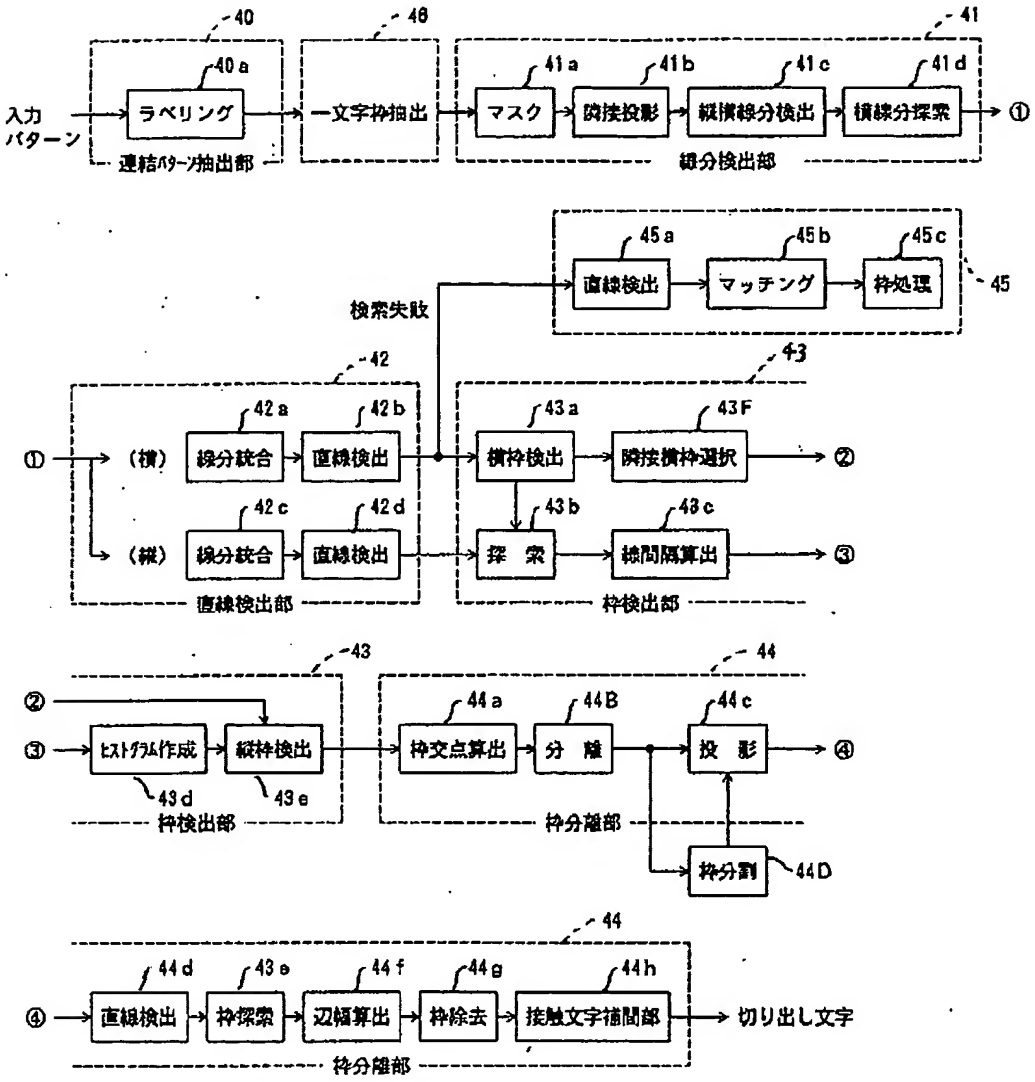

[Drawing 19]

フリーフォーマット枠の分割を説明する図



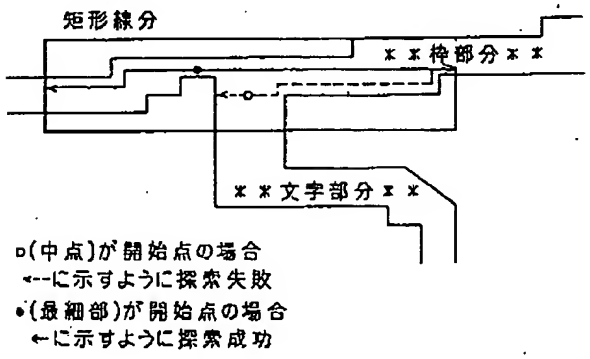

[Drawing 1]

本発明になる画像抽出装置の第１実施例を示すブロック図



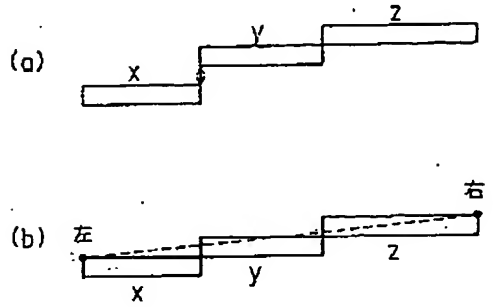

[Drawing 7]

探索の開始点を説明する図



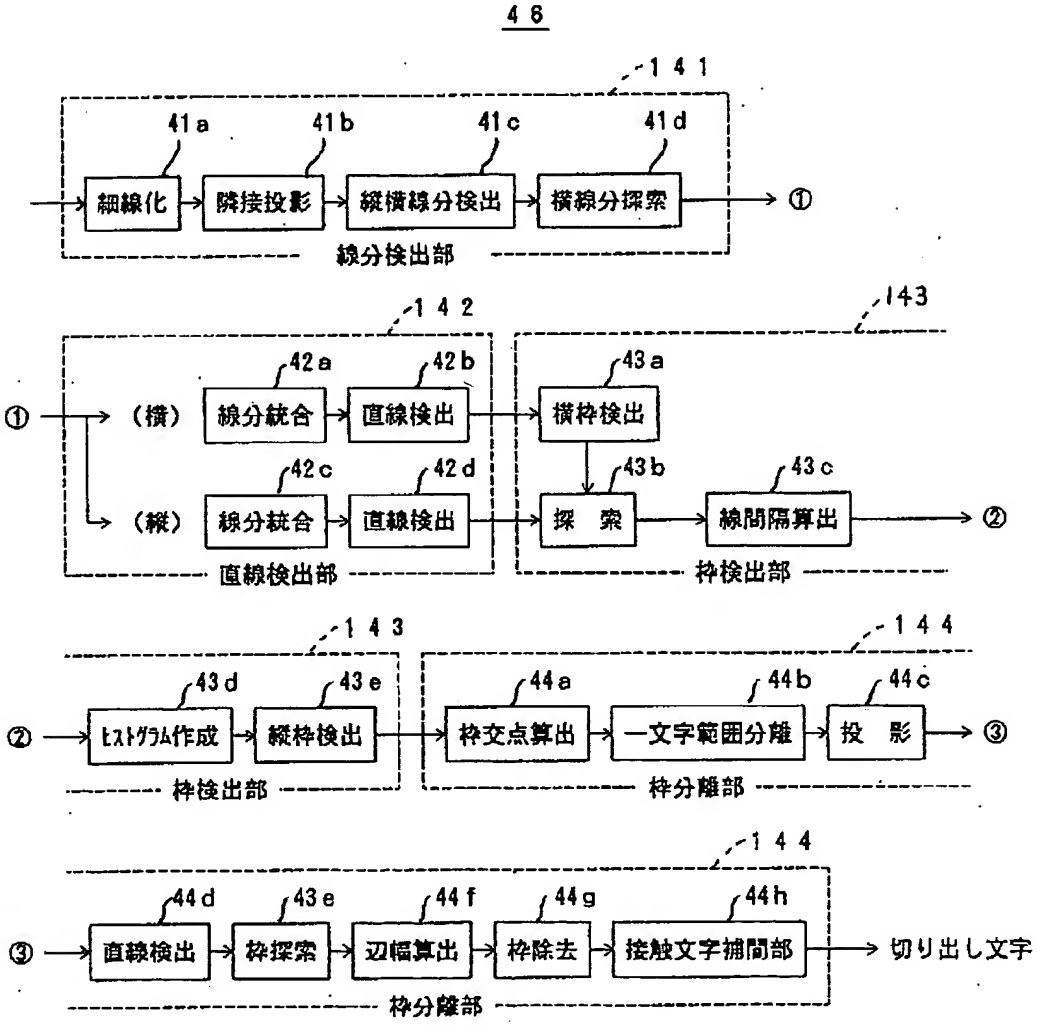

[Drawing 9]

矩形線分の連結及び直線の傾きを説明する図



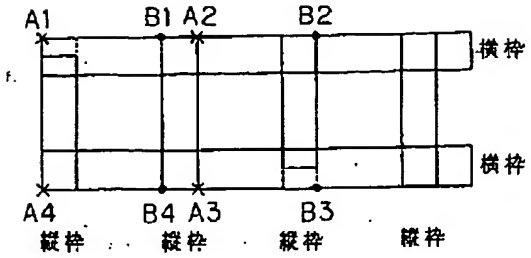

[Drawing 2]

一文字枠抽出部の一実施例を示すブロック図



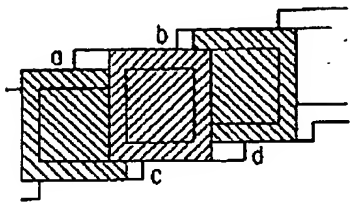

[Drawing 12]

一文字枠への分離を説明する図



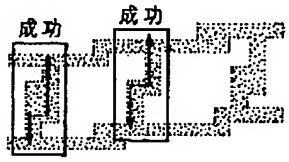

[Drawing 13]

枠の除去範囲を説明する図

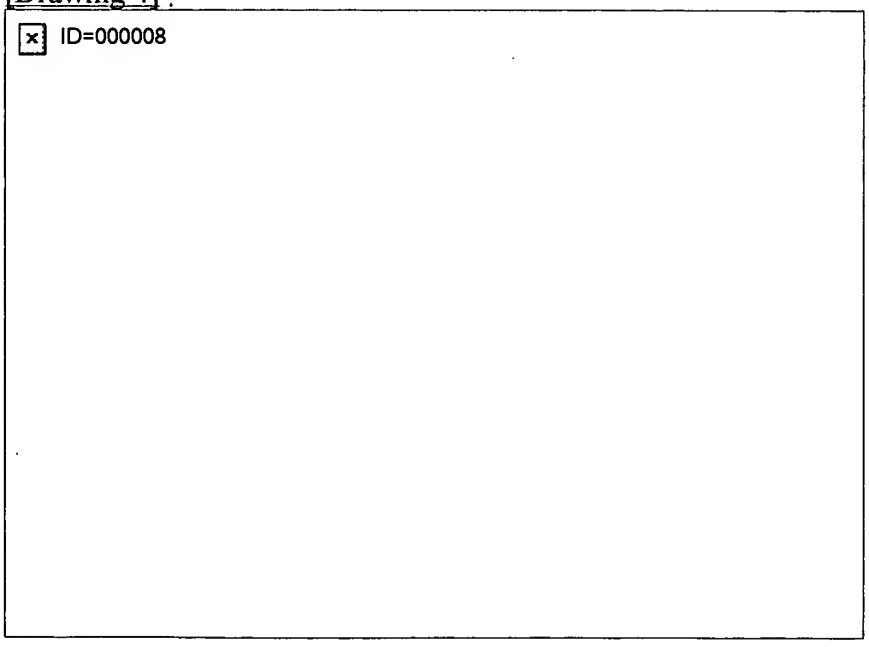



[Drawing 26]

線分の探索を説明する図

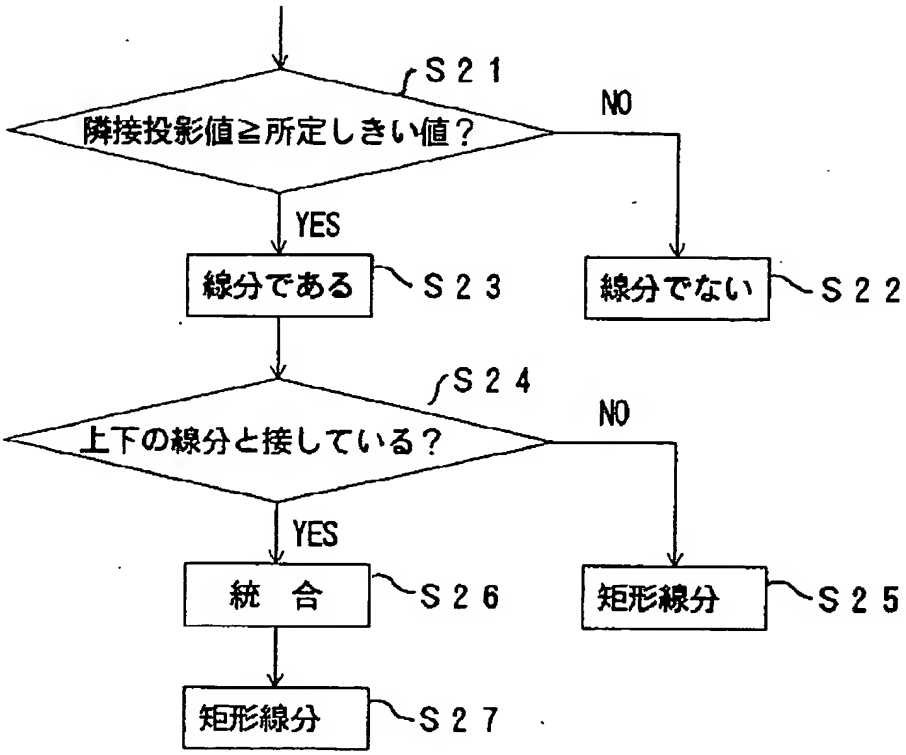



[Drawing 4]



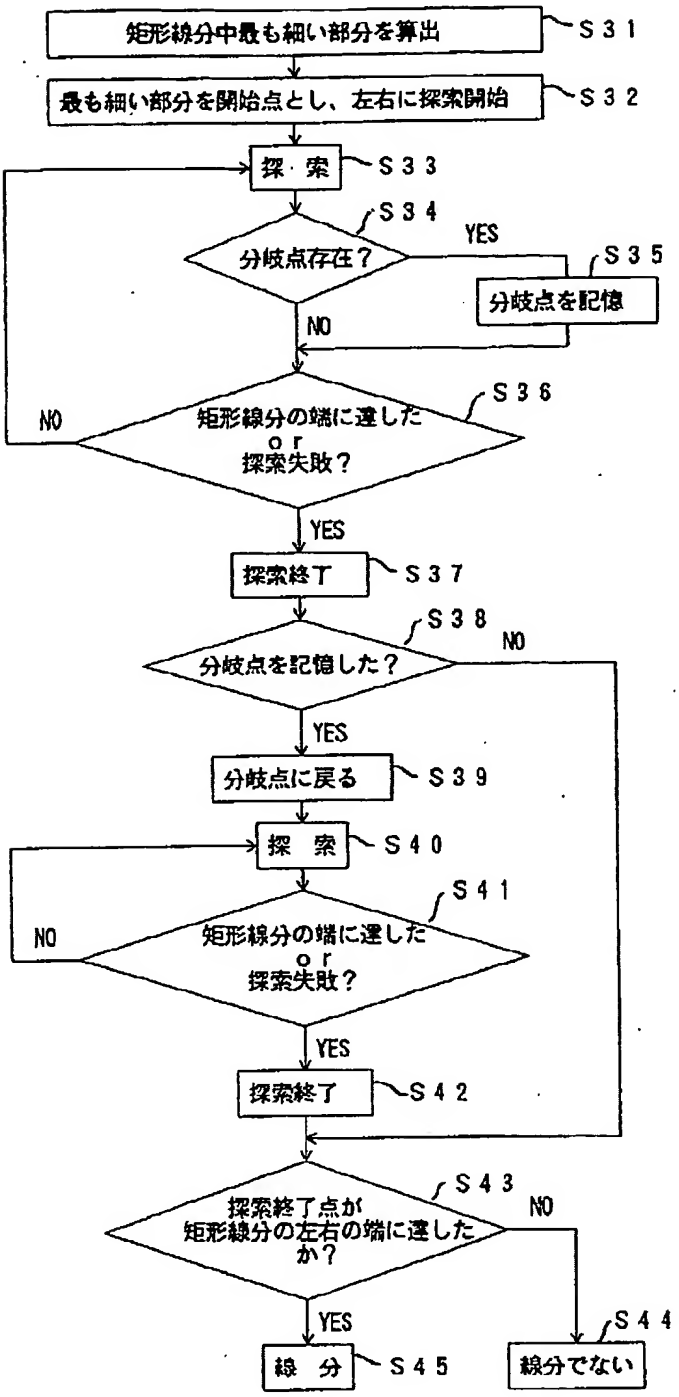

[Drawing 6]

縦横線分検出部４１ｃの処理を説明するフローチャート



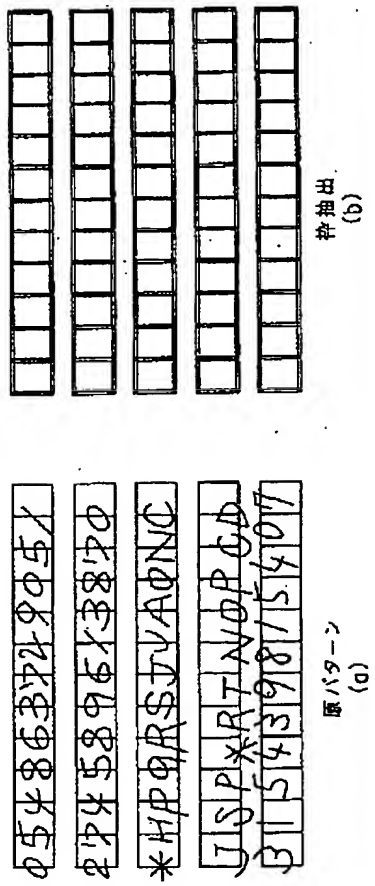

[Drawing 8]

横線分選択部４１ｄの処理を説明するフローチャート



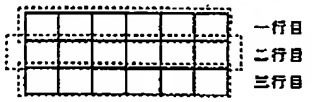

[Drawing 14]

枠の抽出を説明する図

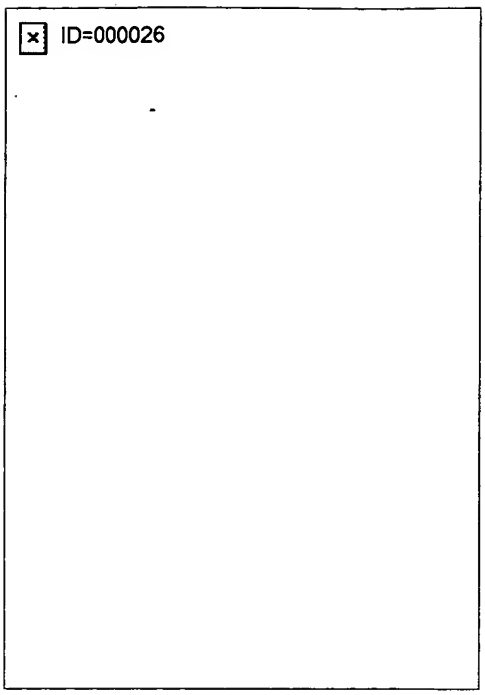



[Drawing 20]

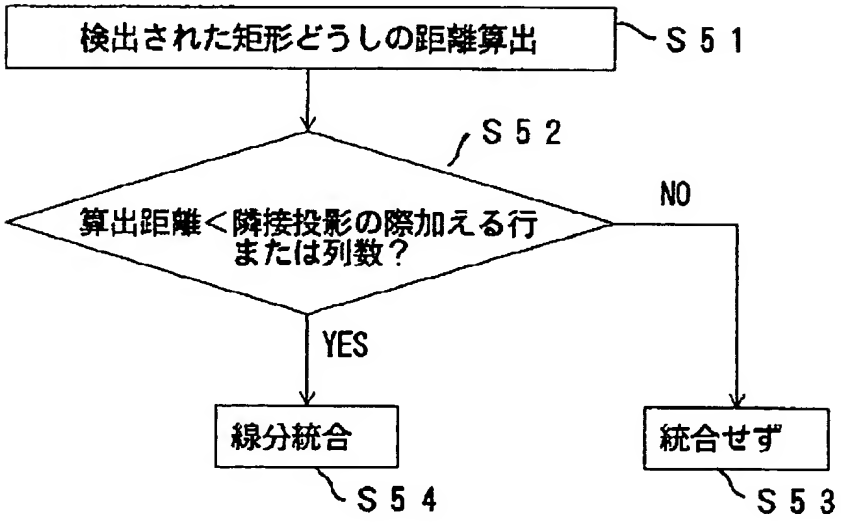
変形式のブロック枠におけるブロック枠の重なりを説明する図




[Drawing 22]



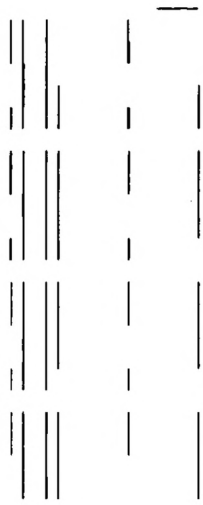

[Drawing 10]
線分統合部４２ａ，４２ｃの処理を説明するフローチャート




[Drawing 15]

文字の切り出し及び補間を説明する図




[Drawing 16]